

Run on: January 17, 2002, 10:21:00
(without alignments)
3020.466 Million cell updates/sec

Title: US-09-740-211-13_COPY_11000_11933
034
cttcaactgacccctgcag 934

perfect score
sequence:

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.00

Searched:	1472140 seqs, 8248589755 residues	2944280
-----------	-----------------------------------	---------

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

```

Post-processing:	Minimum match:	100%
	Maximum Match	100%
	Listing first	45 summaries

Database :

```

1:  gb_da:*
2:  gb_hrg:*
3:  gb_in:*
4:  gb_ov:*
5:  gb_pat:*
6:  gb_ph:*
7:  gb_pl:*
8:  gb_pr:*
9:  gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_pl:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_hrg_hum:*
31: em_hrg_inv:*
32: em_hrg_rod:*
33: em_hrg_hum:*
34: em_hrg_rod:*
35: em_hrg_rod:*
36: em_hrg_other:*

```

30. $EN_{ij} = \frac{1}{N} \sum_{k=1}^N \max(0, \text{pred. No.} - \text{score})$ is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	934	100.0	11933	6	AR138377	AR138377 Sequence
2	934	100.0	11933	6	AR146687	AR146687 Sequence
3	934	98.8	42529	12	CVU339284	U39284 Cloning vec
4	923.2	98.8	42530	12	CVU39285	U39285 Cloning vec
5	923.2	98.8	42531	12	CVU39286	U39286 Cloning vec
6	923.2	98.8	42704	12	CVU37692	J02459 Bacterioph
7	923.2	98.8	48502	7	LAMC6	J02453 Cloning vec
8	921.6	98.7	9170	12	UU02453	U02427 Cloning vec
9	921.6	98.7	9205	12	XU002427	M8294 Bacterioph
10	657.4	70.4	735	7	LAM1PR	M8295 Bacterioph
11	448.2	48.0	751	7	LAMPR	X60309 Coliphage H
12	436.2	46.7	805	7	STHRK220	AE069338 Bacterioph
13	436.2	46.7	40751	7	AE069308	AE069338 Bacterioph
14	429.8	45.7	38297	7	AE069352	AE069352 Bacterioph
15	426.8	45.7	39732	7	AE069329	AE069329 Bacterioph
16	425	45.5	6709	7	BP5011580	AJ011580 Bacterioph
17	409.8	43.9	10390	1	AE005492	AE005492 Escherich
18	409.8	43.9	270365	1	AP002562	AP002562 Escherich
19	401	42.9	5963	7	POE22MIN	X78401 Bacterioph
20	401	42.9	41724	7	AE217253	AE217253 Bacterioph
21	125.8	13.5	191683	2	AE009667	AC009667 Homo sapi
22	58	6.2	11030	1	AE005288	AE005288 Escherich
23	58	6.2	222605	1	AP002555	AP002555 Escherich
24	48.4	5.2	10582	1	AE000253	AE000253 Escherich
25	48.4	5.2	17300	1	D90800	D90800 E.coli geno
26	48.4	5.2	19465	1	D90798	D90798 E.coli geno
27	48.4	5.2	20524	1	D90799	D90799 E.coli geno
28	42.2	4.5	137956	2	AC021065	AC021065 Homo sapi
29	42.2	4.5	179340	2	AC021065	AC021065 Homo sapi
30	42.2	4.5	184452	2	CE12831	M25450 Plasmid Col
31	39	4.2	100	1	CE12831	I66494 Sequence 14
32	37.6	4.0	7218	6	166494	L16621 Caenorhadi
33	37.6	4.0	36977	3	L16621	AC092343 Homo sapi
34	37.2	4.0	178967	2	AC092343	AX024319 Sequence
35	36.6	3.9	37856	1	AX024319	AX024212 Sequence
36	36.6	3.9	37856	6	AP024212	AP023164 Zeta may
37	35.6	3.8	2310	8	AP023164	AP002376 Sequence
38	35.6	3.8	100079	2	TI1G11	TI1G11
39	35.6	3.8	155653	2	AL157390	AL157390 Homo sapi
40	35.6	3.8	163507	2	AL1591134	AL1591134 Homo sapi
41	35.6	3.8	171800	2	AL160060	AL160060 Human DNA
42	35.4	3.8	49054	2	AC019890	AC019890 Drosophill
43	35.4	3.8	163095	2	AP002523	AP002523 Oryza sat
44	35.4	3.8	324484	3	AE003584	AE003584 Drosophill
45	35.2	3.8	893	10	MM0011734	AE011734 Mus muscu

ALIGNMENTS

PAT 16-JUN-2001

RESULT 1	PAT	16-JUN-2001
AR138377		
LOCUS		
AR138377	11933 bp	DNA
DEFINITION	Sequence 13	from patent US 6200560.
ACCESSION	AR138377	
VERSION	AR138377.1	GI:14480722
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	Unclassified.	
ATTORNS	1 (bases 1 to 11933)	
TITLE	Conoulo, P.B., Colosi, P.C. and Qian, X.	
JOURNAL	Adeno-associated virus vectors for expression of factor VIII by	
FEATURES	target cells	
source	patent: US 6200560-A 13 13-MAR-2001;	
	Location/Qualifiers	
	1..11933	
	/organism="unknown"	
BASE COUNT	3258 a	2818 c 2717 g 3140 t
ORIGIN		

LOCUS	AR146887	11933 bp	DNA	PAT	08-AUG-2001
DEFINITION	Sequence 13 from patent US 6221349.				
ACCESSION	AR146887				
VERSION	AR146887.1	GI:15110690			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 11933)				
TITLE	Conto,L.B., Colosi,P.C. and Qian,X.				
JOURNAL	Adeno-associated vectors for expression of factor VIII by target				
FEATURES	cells Patent: US 6221349-A 13 24-Apr-2001.				
source	Location/Qualifiers				
	1..11933				
	/organism="unknown"				
BASE COUNT	3258 a	2818 c	2717 g	3140 t	
ORIGIN					

Query Match	100.0%	Score 934	DB 6	Length 11933
Best Local Similarity	100.0%	Pred. No. 1,5e-266		
Matches 934	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy 1	aaagcgaggaagatctgtatagcgacctctctgcaccgcgcttccagaaggtcatgtgta			
Db 11000	AAAGCGACGAGAGGTGTATAGCGGAGCTCTTGCGAACCCGCTTTCACGAAGGTATGCTTA			
Qy 61	aaagcgagcgagcgtaacctattactaatgaattcagaacagacagltgctacggtcagtt			
Db 11060	AAAGGCGCCACGCGTAACCTATTACTAATGAAATTCAGACGACAGACAGTGGCTAGGCTAGTT			
Qy 121	tgggttgcgcgtgtgctgctggcgagatbaagcgccgttcagatttggatgcatcgctctgc			
Db 11120	TGGGTGTGCTGCTGTGCTGGCGGCGCATGACGCTGTACGCAATTTGGTGTATCCGGTTCGC			
Qy 181	ttccggtattcgttaattcagacaagaagaaagacatgctgaaccagctctgcgcgac			
Db 11180	TTCCGGTATTCGCTTAATTCAGACAAACGGAAGAACACTGTGCTAACCGCTTGCGGAC			
Qy 241	tcttcacgattatgacatcaatgctcttaccctgtgtgcagatataaaaatccgaagc			
Db 11240	TCCTTCACGATTTATGCACATCAATGCTTCTTACTGTTGTGCAGATATATAAAATCCCGAAG			
Qy 301	cgatttcagagcgctcctaactattaccctgcgaactgtttcggaattgcacatttgcagacc			
Db 11300	CGTATTCAGAGGCTCTTAACTATATACCTGCGAACTGTTTGGGATTTGCAATTTTCGACGCTC			
Qy 361	tctgcctgcagatggttggaattccagagcagatagctgcgaatggaaccaatgacgaatc			
Db 11360	TCCTGCTGCATGCTTTGAGATTCCAGACGATACGTCCAACTGACCAACTAGCGGGAATCG			
Qy 421	gtatgaagcgccgcctcttcaatctcaaccacacagcggaattaaacccatcgltga			
Db 11420	GTAATGAGCGCGCCCTCTTTTCATCTCACTACACAAACGAGCAATTAACCCATCTGTGA			
Qy 481	gtcaaatttacccaatttattcaatagccaatatcatgcggttaatatatgtgcacac			
Db 11480	GTCAAATTTACCCAAATTTTATCAATAGTCATATCAATAGCCGTTAATATGTGCCATCC			
Qy 541	gtggcaatcatgctgcgtcaagctgtgcacgcatcacaatgtgtctgcgattgaacctc			
Db 11540	GTGGCAATCATGCTGTACAGCTGTACGCCCATTCAAAATGTGTGCTCGATTACCTCTTC			
Qy 601	tttgtgcgattgcacacacacagagcgatcacaagcggtctaaagctcgtgcaccaggtgg			
Db 11600	TTTGTGCGATTTGACACACACAGCGCTCATACAGCGGCTTAACAGTGTGACCAAGTGGG			
Qy 661	ttgggttaaggtttggatgatgcatgctcaagcgcgatatgctgcgctctgcgagatcct			
Db 11660	TTGGGTATAGGTTTGGATTTAGCATGCTCACAGCGCAATGTGCTGCGCTTGCGCATCTCT			
Qy 721	tgaatagcgagcgcttgcabcttccgcaacttctcagacaactctccccaagctc			


```

|||||
Db 11720 TGAATACCGACGCTTTCGACCTTTCGACAACTCCGCCACAGCTC 11779
OY 781 tctttggaataatcaacgcagcgccttaccatggcaatctcgcattgcccccg 840
Db 11780 TGTATTGGCAATATCAACGCGACGCTTACCATGGCAATCTGCACTTGCCCCGG 11839
OY 841 cgttcgagcaacgagcaataaccgataacgaatgttcgagcaacttgcagtaactt 900
Db 11840 CGTCGCGGCACTACGCGCAATATCGCATTAAGCAATGTTCGAGACACTTGCAGTACCTT 11899
OY 901 tgccttagtattcttcctcaagcgtgcctcgacg 934
Db 11900 TGCCTTAGTATTCTCTTCAAGCTGCCCTGCAGG 11933

RESULT 3
CVU39284/c 42529 bp DNA SYN 13-APR-1996
LOCUS Cloning vector TLF97-1, lambda phage lacZ translational fusion
DEFINITION vector, complete sequence.
ACCESSION U39284.1 GI:1066304
VERSION U39284.1
KEYWORDS Cloning vector TLF97-1.
SOURCE Cloning vector TLF97-1.
ORGANISM artificial sequence: vectors.
REFERENCE 1 (bases 1 to 42529)
AUTHORS St Pierre, R. and Linn, T.
TITLE A refined vector system for the in vitro construction of
a single-copy transcriptional or translational fusions to lacZ
JOURNAL Gene 169 (1), 65-68 (1996)
MEDLINE 96186904
REFERENCE 2 (bases 1 to 42529)
AUTHORS St Pierre, R.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
source
1. 42529
/organism="Cloning vector TLF97-1"
/db_xref="taxon:43840"
complement(19404..19489)
/note="trnB2 terminator"
complement(19507..20760)
/gene="lacY"
complement(19507..20760)
/gene="lacY"
/codon_start=1
/transl_table=11
/product="lactose permease"
/protein_id="AAC53648.1"
/db_xref="GI:1066305"
/transl_table="MYILKNTNEMWGLFFFEYFIMGAFPEPPTMLDINHISKSD
TGIIFATSLPFLILFOPLGLISDKLRLKLYLMTITGMVMAPEFIFEGPLQYN
ILVSTYIGITLGCFCNAGAPAYAEATIEKYSRSNPEFGARMEGVGALISLGI
MTINNOVFWLGSICALILAVLFFKTKTAPSSATYANAVGNHSAFSLKIALFIR
QPKLWFLSLVIGVSCYDVDFDOQFANFPTSFRTGQGRVFGYTTMKELLNLSIM
FFAPLILIRIGKNAALLAGTAKMSVRIGSSFAISALEVILTLHFEVPLLVGCF
KYTSQFEVRSATYLVCFEKFOLAMIFMSYLAGNMEISIGFQAGYIVLGLALGE
TLISYFTLSGPPSLLRQYNEVA"
complement(20812..23859)
/gene="lacZ"
complement(20812..23859)
/note="EcoRI site normally present at the 3' end of lacZ
has been removed by an uncharacterized mutation"
/codon_start=1
/transl_table=11
/product="beta galactosidase"
/protein_id="AAC53649.1"
/db_xref="GI:1066306"
/transl_table="VVIQORRDWENPGVTLNRLAHPFASWNSSEARIDRSQQLR

```

```

SLNGEMRFPAPPAPEAVESWLECDLPEADTVVVPNSMNGHYDAPITYNVITVN
PPEVTEPNTQCYSLTEPVENSSMLDEQRTIIFGVNVAEPHLCNGRWVYGQDSRLP
SEEDLSADFLRAGENELAMVLRMSDGLSTIEDQMMRMSGILFRVYSLHKPTQISPH
VATRPNDSPFSAVLEAVQWCGELLDLTARTYSLMOEPTOVAQGTAPGGEIIDENG
VADRVTLLINVENPRLMSAELPNLYRAVELHLDLITAEACDNGRRVRLIENELL
LNGKPLILRGVNRHEHPLHGOVMDQETQVODILLKONKNFAVRSHYPNPLWYT
LCDRYGLVVDENAIETGHPVPMRLDDPFWNTPAMEERYLRMYQRNRNPSYIWSL
GNESGHANHDALYEMIKSDPSRPVQYEGGADTATDIIICPMYARVDEDPFVAV
KMSIKKKSLIPEETPRLICEYAHAMGSLGFAKWOAFROYPRLOGGVYMWVDS
LKIYDENPMNATAGDPRDPRNDRCMNGLVFADRPHPALTEAKHQOQFQFRLS
GOTIIVTSYLFHNSDNLHMYALDCKPLASEGEVPLDAPGKOLIELPELPQES
ACQIMLIVAVVOPNATVMSSEAGHISAMQOMLAENLSYTLPAASHAPHTTSEMPC
IELGMRKMOFNRSGFSLQMWIGDKOLITPLRQDFTRAPLNDIGVSEATRIDPNM
VERKRAAGHYOAEALLQCTADTLADAVLITTAHAWHQGKTLFISRTYRIDSGOM
ATTVCEVAVASDRPHPARIGLNCOLAQVAREVNMGLGFOENYPRDLTAACFDRWDLPL
SMYTPYVPSRNGICRGTRNLNGPROMDGFQENISRYSOQOLMETSHRHLHABE
GTWLNIDGFHMGITGDDSMSPSVSNPRLDSAGRYHYOLWYQK"
complement(23874..23957)
/note="multiple cloning site"
complement(23973..24141)
/note="trnB1 terminator"
24475..25665
/gene="tetR"
24475..25665
/gene="tetR"
/codon_start=1
/transl_table=11
/product="tetracycline resistance protein"
/protein_id="AAC53650.1"
/db_xref="GI:1066307"
/transl_table="MKSNNALIVLIGTVTLDAVGLVLPVPLGLLRIVHSDTASH
YGVLLALVLMQFLCAPVGLALSDRGRPVLLASLDATTDYAIAMATTPVMTIYAG
RIVAGITGATGAVAGAYIADITDGEDRAHFEIAMSACGVCAVGPVAGVGLGALISLH
APPLAAVINGNLILGCEFLMOESHGERRPMPPLPAFNVSSEFRMARMTIYAALMTV
PTIMOJVGOPBALWYIFGEDRRMSATWIGSLAVFGILHLAQAFTGATKTFGE
KQATILQMAADALGYVLLAFATPGMAAPFIMILLASGIGMPALDAMLSROVDDHOG
OIGSLAALVLSITGPIVTAIYTAASSTWNGLAMIVGALYLVLCPLBRGAWSR
ATST"
33602..35450
/note="hmm21 substitution"
BASE COUNT 10291 a 10533 c 11616 g 10081 t 8 others
ORIGIN
Query Match 98.8%; Score 923.2; DB 12; Length 42529;
Best Local Similarity 99.7%; Pred. No. 4,4e-283;
Matches 925; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 aacgcgcagagaggttqtagcggaactcctccgacccgccttcacgaaggtcattgta 60
Db 39091 AAACGCGAGAGAGTGTAGCGCGACCTCTGCCACCCGCTTTCAGCAAGGTCATGTGTA 39032
OY 61 aaagcgcgcagaggttaccattaccattacgaatcaggaacagtggtcaggtcagtt 120
Db 39031 AAAGCGCGAGGCTACTATTACTATTAATTCAGACAGACATGCTACGGCTCAGTT 38972
OY 121 tgggttgctgtctgtcgtggcgagcagtcagtcagtcagtcagtcagtcagtcagtcagtc 180
Db 38971 TGGGTGTGCTGTGCTGTGCGCGGCGATGACCGCTGTGACGATTTGGTGATCCGGTTCTGC 38912
OY 181 ttcgcgattcgcattcagcacaacggaagagacatggcctcaacacagctcgccagc 240
Db 38911 TTCCGGTATTTCCTTAATTCACACACAGGAAGAGCACTGGCTACACAGGCTGCCGAC 38852
OY 241 tcttcagattcagcacaacgctcttaccctgtgtgtgcagataaanaaaatcccgaaac 300
Db 38851 TCTTTCAGATTATTCAGCTCAATGCTCTTAACGCTGTGTGACAGTATAAAAAATCCCGAAAC 38792
OY 301 cgtatgcaggtcctaactatcactgtggaactgtttgggattgacatttggcagacac 360
Db 38791 CGTTATGACAGGCGCTCAATTAACCTGCGAAGCTGTTTGGGATTCATTTTGCACAGACCTC 38732
OY 361 tctgcctgcagatgttgaggctcccaagacatacgtcgaagtgaaccaactaagcggaatcg 420

```

```

Db 38731 TCTGCCGCGATGTTGGATTCACAGCATACGTGAAGTACCACTAGCGGAATCG 38672
Qy 421 gtgttaagccgcgcctctttatctcactaccacaagcgaaatcaaccatcgttga 480
Db 38671 GTGTGATGCGCCCGCTTTTATCTCATCTACCAACAGGGAATTAACCATGTTGA 38612
Qy 481 gtcaaatiaaccaattatcaataaagatcaatcagcgttaataatgttccatcc 540
Db 38611 GTCGAATTTACCAATTTATTCATTAAGTCAATTCATGCCGTTAATATGTTGCATCC 38552
Qy 541 gtggcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 600
Db 38551 GTGGCAATCATCTGCTTACAGTGTACCGCATTTAAATGTTGCTCGATTTGACTCTTC 38492
Qy 601 ttgtgtgcatcgcacacacagagcgatcaacagcggtcgtcgtcgtcgtcgtcgtc 660
Db 38491 TTTGTGCGATTGTGACACCGAGCGCTCATACGCGCTTAACAGTGTGCGTGAAGTGG 38432
Qy 661 ttgggttaaggtttgggattagatcgtacacagcgcgatatactcgtcgtcgtcgtc 720
Db 38431 TTGGGTAAAGTTTGGGATTAGATGTCACACCGGATATCTGCGTGTGCGATCTCT 38372
Qy 721 tgaatagccgaagccttgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 780
Db 38371 TGAATAGCGAGCGCTTGTGACCTTCCGACCTTTCTCGACACTCTCCCGCACACTC 38312
Qy 781 ttttttgcaatatacaacacagcgagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 840
Db 38311 TGTTTTGGCAATATCAACCGACGCGCTGTACATGGAATCTGTGATCTTGTGCGCCGG 38252
Qy 841 cgtcgcgcactaagcgcaataatccgcaataagcgtcgtcgtcgtcgtcgtcgtcgtc 900
Db 38251 CGTCCGCGACTACGCGCAATATCCGCATTAAGCAATGTTGCGAGCACTTGCAGTACTC 38192
Qy 901 tgcctagatcttcctcctaagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 928
Db 38191 TGCCTACTATTCTTCTTAACCTTTGCC 38164

RESULT 4
LOCUS CVU39285 42530 bp DNA SYN 13-APR-1996
DEFINITION Cloning vector TLF97-2, phage lambda lacZ translational fusion
ACCESSION U39285
VERSION U39285.1 GI:1066308
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
terminator
gene
CDS

```

Cloning vector TLF97-2.
Cloning vector TLF97-2.
artificial sequence; vectors.
1 (bases 1 to 42530)
St Pierre, R. and Lim, T.
A refined vector system for the in vitro construction of
single-copy transcriptional or translational fusions to lacZ
Gene 169 (1), 65-68 (1996)
2 (bases 1 to 42530)
St Pierre, R.
Direct Substitution
Submitted (24-OCT-1995) Thomas Lim, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6A5S1, Canada
Location/Qualifiers
1. 42530
/organism="Cloning vector TLF97-2"
/db_xref="taxon:43841"
/complement(19404..19489)
/note="rnmB2 terminator"
/complement(19507..20760)
/gene="lacy"
complement(19507..20760)
/gene="lacy"
/codon_start=1

misc_feature
BASP COUNT 10292 a 10533 c 11615 g 10082 t 8 others
ORIGIN

Query Match 98.8%; Score 923.2; DB 12; Length 42530;
Best Local Similarity 99.7%; Pred. No. 4.4e-283;
Matches 925; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 aaacggcaggaaggttgtagcgcgacccctcgcaccccgcttcacgaagtcgtgta 60

Accession	Source	Cloning vector	TF97-3.
39092	AAACGGACGAGGTTGTTAGCGGACCTCTCCACCGCGCTTACACAGTGCATCGTGA		39033
61	aaagccgcagcgttaactatataatgaatcaagaagaacagtgtaacgctcaagtt		120
39032	AAAGCGCCGACGGTAACATTAATGAATTCAGACAGACAGAGGGGTACGAGCTCAGTT		38973
121	tgggttgctgtctgtcgtgcgagcgatgacgctctgaacgcttggatgcggtctgc		180
38972	TGGGTTGCTGTCTTCTGCGGCGGATGAGCGCCGTGACGATTTGTGATCCGGTCTGC		38913
181	ttccgylatcgcttaatttcagacaagaagagagactgctgaaccaagctgcgcac		240
38912	TTCCGGATATCGCTTAATTCACACACACAGGAAAGACACGCTTAACCGAGCTGCCAC		38853
241	tcttcacgattacgcactcaatgctcctactctgtgtgcgatalablaaaaaatccgaaac		300
38852	TCTTCACGATTATTCGACTCAATGCTCTTAACGTCGTGACGATATATAAAATCCCAAC		38793
301	cgtatcgcagcgtcttaactataccgcggaactgcttgcgatttgatcttgcagctc		360
38792	CGTTATGACGCTCTTAACCTTACCTGCCAACCTGTTTGGGATGGAATTTTGGACAGCTC		38733
361	tctgcctcgcgtatgctgcgaggttcgaagcagatgcgcgaagtgacaactagcggaatcg		420
38732	TCTGCCTCGATAGGTGTGAGTTCACAGACATATGCTGGAATGATCAACACTAGCGGAATCG		38673
421	gtatgaagcgcgcctcttctatcctcaactacaagaagagaattaacaccatgcttga		480
38672	GTATGATGCGCGCCTCTTTCATCTCATACACACACAGCGAATTAACCATTCGTTGA		38613
481	gtcaaatcccaatttatcatcaataagtaataatcatagcgttaataatgctgcatcc		540
38612	GTCAAAATTTTACCCAAATTTTATTCATTAAGTCAATATCATCCGTTATATGTGCGATCC		38553
541	gtgcacatcatctgctctaaacgtgtgacgcgcctcaaaatgttgtctgcgattgactctc		600
38552	GTGCGCATCATCGCTGTAAAGTGCTGACGCAATTCAAAATGTGCTGCGATTCGATACCTTC		38493
601	tttgtgagattgacacacacacagagcgtcatcaagcgtcttaacagtgctgcagacagttgg		660
38492	TTTTGTGGCATTTGACACACACAGACGCTCATACGCGGCTTAAACAGTGGCTGACACAGGGG		38433
661	ttaggttaagtttgtgagttgcatctgcaacagcgatagcgtgacgtgctgtctgcatcct		720
38432	TTTGGTAAAGGTTTGGATTATACATCGTCACAGCGCATATGCTGCGCTGTGGCATCTCT		38372
721	tgaatagcgcagcctttgcaatctccgcacctttctggaacctctcccccagaqtc		780
38372	TGAATAGCGGACGCTTGTGCATCTTCGCACTCTTCTGACACACTCCGCCACAGCTC		38312
781	tgatttggaalatacaacgcagcgcgtgtaccaatgagcaatctctgagatctgtcccccgg		840
38312	TGTTTTGGCAATATCAACGCGACGCGCTGTACCAATGGCAATATCTGCATCTTGCCCGCG		38252
841	cgtgcgcgcactacgcgcaataatgcgcaataagcgaatgltgacagacttgcaagactt		900
38252	CGTCGCGGCACTACGCGCAATATCCGATTAAGCGAATGTTCGAGCACTTGCAATCTT		38192
901	tgactagatattctctcaagctgcgcc	928	
38192	TGCCTTAGTATTTCTTCAAGCTTGGCC	38165	
RESULT 5			
CVU39286/c	42531 bp	DNA	SYN
LOCUS	CVU39286	Cloning vector TLF97-3, phage lambda lacZ translational fusion	13-APR-1996
DEFINITION	vector, complete sequence.		
ACCESSION	U39286		
VERSION	U39286.1	GI:1066312	
KEYWORDS			

ORGANISM	Cloning Vector (LacZ)	artificial sequence; V
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

REFERENCE
AUTHORS

1. (bases 1 to 42524)
St. pierre, R. and Linn, T.

TITLE	A retained vector of single-copy transcripts

Gene 169 (1), 96186904
JOURNAL
MEDICINE

REFERENCE
AUTHORS
2 (bases 1 to 42000)
StPierre, R.

TITLE	Direct submission
JOURNAL	Submitted (24-OCT-1995)

[illegible]

```
source
1.42001
/organism="Clo
```

```
complement(194
/ab_x1e1
terminator
```

complement (19)

complement (19
"1357"
/gene
CDS

```
/codon_start=
```

```

/product="1acc

```

```
/db_xref="GI:
```

7
TGI FAISLESI
TUSCIVGIVIA

MEETINQ EVFWLQ
COTWEL SI VVITQ

FEAPLINRIGG
ZVTSCOEVBES

TLISVFTLSGPG
complement (2

```

    /gene="lacZ"
    complement(2

```

```

cds
/gene="lacZ"
/nota="ECORI"

```

has been rem
/odon start

```

./transl_tabl
/product="be

```

protein_id=

/translation
STINGEWFAME

PFVPTENPTG
SEFDLSAFLRA

VATRENDDESK
YADRVTLRLNV

LNKGKPLLIR
ICDRYGLYVL

GNESGCHGANTL
KWSIKKWLSE

LIKJ DENG
GOTIEVISEYI

AGQLWLI V K
IELGNKRWQFI

VERWKAAGH
AITVDYEVAS

SDMLIFVH
GTWLNIDGFH

```
misc_feature
```

```
terminator
completion
/note="rm
35
```

gene

CDS
247 / 1 / 1
/gene="tel

1


```
/transl_table=11
/product="beta-galactosidase"
/protein_id="AAC33646.1"
/db_xref="GI:1051183"
/translation="MTMTIDSLAVVLRDRDMPGVOTUNRLAAHPFASMRNSPEAR
TDRPSQULSLNGEWRFAFPAEAVPSPMSLECDLPADIVYVPSNMOMHGYDAIY
NTVYITVNPVPVPEPPEPCCYSLFENDESLODGCORILFDGYSNPFHLMGCRVY
GYGDSRLPSEFEDLSAPLRAGFNRLAVVLRNSDSYLDDOMMMSISIFRVSLLHK
PTTQISDHFVATRFNDPFSRAVLEAVYVMSGELRYLSDYVLSMOGTEVAGCPFG
GFIIDERGQYADRVTLRLNVENPKLMSAEIPLYAVAVELTADTLEIENACGVAFG
EVRINDEGLLLNGKPLIRGVNRHEHHLHCQVMDQVODIILMKNNANVACSH
YFNHLYWTLICDRGYLYVDEANIEFHGMVPMNRITDPPHMLPAMSEVETVMQDRN
HPSVLIWSIGNSGSHANHDAIYRMIKSVDSRPVQYGGCAADTATITITCPMYARD
EDPPAVPKSKIKKSLSPGTRPLILCEVRAHANGSLGFAKYMAFROYPRLOGS
FYMPVVDOSLIKIDENGNPWSYGGDFDTPMDRQFCNMGVLVADRPPHAPALTEAKH
QOFPRLSGGTEVTSYLRHSDNELHMMVAILDGPASGEVPLDVAOQGLIE
LPFLPQSPASQOLMLTVQVQNPATVMSQWISMOQWRIENSLVTLPAASHAPIH
LTTSEMPICIELGNKRMQFNOSGFLSQMTIGDKKOLLTPRDOETRAPLDNDIGVSE
ATRIDPNAMVERMKAAHQVQAEALLOCTADTLDVAVLITTAHMOGKTLFISRTT
YRIDSGDMAITDVEVADTPHPARISGLNOIAQVARYWMLGIGQDENPDLRTAA
CEDRMDPLSDMYTPYVPSSENGLRGCRRELNTNGPHQWGDFOFNISIRY500LMTS
HRHLHAEGTWINIDGTHMGIIGDSDSPSVSAEFLSAGRYHOLWQCR"
complement(24019..24132)
/feature
misc_feature
/note="multiple cloning site"
complement(24148..24316)
/note="rrnB1 terminator"
24650..25840
/gene="tetR"
24650..25840
/gene="tetR"
/note="tetR"
/note="Description: tetracycline resistance gene;
tetracycline resistance protein"
/codon_start=1
/transl_table=11
/protein_id="AAC33647.1"
/db_xref="GI:1051184"
/translation="MKSNMALVILGTVTLDAVIGLVMPLPGLIHDIVHSDIASH
YGVLLAYALMQLCAPVLGALSDRFRRLPLASLIGATIDVAIMATVTLVLYG
RIVAGITGATGAVAGAYIADITDGEDRARRHLGMSACFGVMVAPAGGLILASIS
APFLAAVLNGLNLGCFLEMSHKEGRRPMRANPVSPPMAGMTIVAAIMY
FEIMQIAGVPAALMVIIEGDRPRMSATMGISLAVGIIHLALQAVTPYKTRDE
KQAIAGMADALCYVILATRTGMAFPIMILLASGSIQMPALQAMLSQVDDHOG
QLOGSLAIVLSVITGPIVLITAIYASASTWGLAMIVGAALYVCLPALRGAMSR
ATST"
misc_feature
31747
/note="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
misc_feature
33777..35625
/note="imm21 substitution"
39188
/note="EcoRI site normally present at this site has been
removed by an uncharacterized mutation"
BASE COUNT 10338 a 10570 c 11656 g 10132 t 8 others
ORIGIN
```

```
Query Match 98.8%; Score 923.2; DB 12; Length 42704;
Best Local Similarity 99.7%; Pred. No. 4.4e-283;
Matches 925; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 aaagcgagaggtgtttagcgagacctctgcacccgcttcaagaagtcagtgtga 60
Db 39266 AAACGGAGAGGAGTGTGTTAGCGGACCTCTCCGACCCGCTTTACGAAGGTATGTGTA 39207
Qy 61 aaagcgagaggtgtttagcgagacctctgcacccgcttcaagaagtcagtgtga 120
Db 39206 AAAGGCCGACGCTAATTACTAATGAAATTCAGACAGCACTGCTACGGCTCAGTT 39147
Qy 121 tgggttgccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 180
Db 39146 TGGGTGTGCTGTTGCTGGGGGCGGATGACGCTGTACGATTTGGATCCGGTCTC 39087
Qy 181 ttccggtatcgccttaattcaagcacaagagagcactgcgtcaaccagctcgcgcac 240
```

```
|||||
Db 39086 TTCCGGTATTCGCTTAATTCAGACACAGCAAGAAAGACTGGCTAACAGGCTGCCGAC 39027
Qy 241 tcttacgattatgcaccatgccttactctgttgcagatatataaaatccgaac 300
Db 39026 TCTTACGATTATGCACATGCTTACCTTACCTGTTGTCACATATATAAAAAATCCCGAAC 38967
Qy 301 cgtatgcagctccttaattacatcgcagactgttcggattgcatcttgcagactc 360
Db 38966 CGTATGACAGCTCTAATCTATTACTGGGAAGTGTTCGGATTGCAATTTGCGAGACTC 38907
Qy 361 tctgcctgcagatggttgcagttcagacagatcagtcgaagtgcacaaactaagcgaaatcg 420
Db 38906 TCTGCTTCGATGATGTTGAGTTCACAGACATATCGTCAAGATGACCAACTAGCGGAAATCG 38847
Qy 421 gtagaagcgcgcgcctcttcatcctaactcacaagcagcgagattaccacatcgltga 480
Db 38846 GTAGTAAGCGCGGCTCTTTTCACTCTCACTACCAACAGACGAGCAATTAACCATCGTTGA 38787
Qy 481 gtcaatttaccatatttattcaataagtcataatcatgcgttaatatgttcacatcc 540
Db 38786 GTCAATTTACCAATTTTATTAATTAAGTCAATATCATGCGCTTAATATGTTGCCATCC 38727
Qy 541 gtgcacatcatgcctgcctcagatcagtgagccgcatcaaaatgttgcgcagatgccttc 600
Db 38726 GTGCACATCATGCTGCTTACCTGTGACCGCATTCGCAAAATGTTGCTCGATTGACTTTC 38667
Qy 601 ttctgcgcatlgtcacccacagacgcatcatcagcgcgttacaagtgctgacagagtgagg 660
Db 38666 TTTGCTGCAATTTGCACACACAGAGGCTATACAGGCGCTTAAACATGCGGTACCAAGGTGGG 38607
Qy 661 ttgggttaaggtttggattagcatcgtacacagcgcgatatcgtcgttgcgttcacatc 720
Db 38606 TTGGGTAAAGTTTGGAATTTAGCATCTGACACAGCGCATGATGCTGCGCTTGTGCAATCCT 38547
Qy 721 tgaatgcagcagccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 780
Db 38546 TGAATACCGCAGCGCTTGTGATCTTCGCACTCTTCTGCAACATCTCTCCCAAGACTC 38487
Qy 781 tgtttgcgaatalcaacgcgcagcgcctgtacatcgtgcaatcctctgcattcgtcccgag 840
Db 38486 TGTTTGCAATATCAACACCGCAGCGCTGTACCATGTGCAATCTCTGCATCTTCCCGCG 38427
Qy 841 cgtcgcgcagcctaagcgaataatccgataagcgaatggttgcgacacttgcagtcactc 900
Db 38426 CGTGCAGCGCACTGAGCAATATCCGATTAAGCAATGTTGCGAGCACTTGAGTACTT 38367
Qy 901 tgccttagtatcttccttcaagctgcgcc 928
Db 38366 TGCCTTAGTATTTCTTCAACAGCTTTGCC 38339
```

```
RESULT 7
LACMG/c LACMG 48502 bp DNA circular PHG 31-OCT-2000
LOCUS Bacteriophage lambda, complete genome.
DEFINITION J02459 M17233 M24325 V00636 X00906
ACCESSION J02459.1 GI:215104
VERSION
KEYWORDS DNA-binding protein; circular; coat protein; complete genome;
origin of replication; repressor; unidentified reading frame.
SOURCE bacteriophage lambda.
ORGANISM Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
Lambda phage group.
REFERENCE 1 (bases 44388 to 44780)
AUTHORS Lebowitz, P., Weissman, S.M., and Redding, C.M.
TITLE Nucleotide sequence of a ribonucleic acid transcribed in vitro from
lambda phage deoxyribonucleic acid
JOURNAL The Journal of biological chemistry. 246 (16), 5120-5139 (1971)
MEDLINE 71288594
PUBMED 4936723
REFERENCE 2 (bases 1 to 12)
AUTHORS Wu, R. and Taylor, E.
```

- TITLE Nucleotide sequence analysis of DNA. II. Complete nucleotide sequence of the cohesive ends of bacteriophage lambda DNA
JOURNAL 71209066
MEDLINE 4931680
REFERENCE 3 (bases 45493 to 45963)
AUTHORS Imada, M. and Tsugita, A.
TITLE Amino acid sequence of lambda phage endolysin
JOURNAL Nature New Biol. 233, 230-231 (1971)
REFERENCE 4 (sites)
AUTHORS Weigel, P.H., Englund, P.T., Murray, K. and Old, R.W.
TITLE The 3'-terminal nucleotide sequences of bacteriophage lambda DNA
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 70 (4), 1151-1155 (1973)
MEDLINE 73215915
REFERENCE 5 (bases 38597 to 38672)
AUTHORS Dahlberg, J.E. and Blattner, F.R.
TITLE In vitro transcription products of lambda DNA: Nucleotide sequences and regulatory sites
JOURNAL (in) Fox, C.F. and Robinson, W.S. (Eds.), VIRUS RESEARCH. PROCEEDINGS OF 1973 ICA-UCIA SYMPOSIUM: 533-544; Academic Press, New York (1973)
JOURNAL 6 (bases 37945 to 38027)
AUTHORS Maniatis, T., Plashine, M., Backman, K., Kieid, D., Flashman, S., Jeffrey, A. and Maurer, R.
TITLE Recognition sequences of repressor and polymerase in the operators of bacteriophage lambda
JOURNAL Cell. 5 (2), 109-113 (1975)
MEDLINE 75185528
REFERENCE 7 (bases 35583 to 35600)
AUTHORS Kieid, D.G., Agarwal, K.L. and Khorana, H.G.
TITLE The nucleotide sequence in the promoter region of the gene N in bacteriophage lambda
JOURNAL The Journal of biological chemistry. 250 (14), 5574-5582 (1975)
MEDLINE 75189495
REFERENCE 8 (bases 35434 to 35618)
AUTHORS Dahlberg, J.E. and Blattner, F.R.
TITLE Sequence of the promoter-operator proximal region of the major leftward RNA of bacteriophage lambda
JOURNAL Nucleic acids research. 2 (9), 1441-1458 (1975)
MEDLINE 76031664
REFERENCE 9 (bases 37945 to 38018)
AUTHORS Maniatis, T., Jeffrey, A. and Kieid, D.G.
TITLE Nucleotide sequence of the rightward operator of phage lambda
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 72 (3), 1184-1188 (1975)
MEDLINE 75158212
REFERENCE 10 (bases 44588 to 44773)
AUTHORS Sklar, J., Yot, P. and Weissman, S.M.
TITLE Determination of genes, restriction sites, and DNA sequences surrounding the 6S RNA template of bacteriophage lambda
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 72 (5), 1817-1821 (1975)
MEDLINE 75217847
REFERENCE 11 (bases 37905 to 37989)
AUTHORS Walz, A., Pirodda, V. and Inelchen, K.
TITLE Lambda repressor regulates the switch between PR and P_{RM} promoters
JOURNAL Nature. 262 (5570), 665-669 (1976)
MEDLINE 76267718
REFERENCE 12 (bases 37946 to 38039)
AUTHORS Smith, G.R., Eisen, H., Reichardt, L. and Hedegpeth, J.
TITLE Deletions of lambda phage locating a prim mutation within the rightward operator
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 73 (3), 712-716 (1976)
MEDLINE 76152323
- PUBMED 1062780
REFERENCE 13 (bases 35578 to 35667; 37903 to 38027)
AUTHORS Plashine, M., Backman, K., Humayun, M.Z., Jeffrey, A., Maurer, R., Meyer, B. and Sauer, R.T.
TITLE Autoregulation and function of a repressor in bacteriophage lambda
JOURNAL Science. 194 (4261), 156-161 (1976)
MEDLINE 76271154
REFERENCE 14 (bases 35578 to 35667)
AUTHORS Humayun, Z., Jeffrey, A. and Plashine, M.
TITLE Completed DNA sequences and organization of repressor-binding sites in the operators of phage lambda
JOURNAL Journal of molecular biology. 112 (2), 265-277 (1977)
MEDLINE 77209970
PUBMED 875019
REFERENCE 15 (bases 38610 to 38732)
AUTHORS Scherer, G., Hobom, G. and Kossel, H.
TITLE DNA base sequence of the po promoter region of phage lambda
JOURNAL Nature. 265 (5590), 117-121 (1977)
MEDLINE 77100320
PUBMED 834253
REFERENCE 16 (bases 38041 to 38241)
AUTHORS Roberts, T.M., Shimatake, H., Brady, C. and Rosenberg, M.
TITLE Sequence of the *cro* gene of bacteriophage lambda
JOURNAL Nature. 270 (5634), 274-275 (1977)
MEDLINE 78071724
PUBMED 593347
REFERENCE 17 (bases 27616 to 28935)
AUTHORS Davies, R.W., Schreier, P.H. and Buchel, D.E.
TITLE Nucleotide sequence of the attachment site of coliphage lambda
JOURNAL Nature. 270 (5639), 757-760 (1977)
MEDLINE 78071823
PUBMED 593399
REFERENCE 18 (bases 37206 to 37263; 37914 to 37970)
AUTHORS Humayun, Z.
TITLE DNA sequence at the end of the *CI* gene in bacteriophage lambda
JOURNAL Nucleic acids research. 4 (7), 2137-2143 (1977)
MEDLINE 78011659
PUBMED 909767
REFERENCE 19 (bases 27617 to 27934)
AUTHORS Landy, A. and Ross, W.
TITLE Viral integration and excision: structure of the lambda att sites
JOURNAL Science. 197 (4309), 1147-1160 (1977)
MEDLINE 77258934
PUBMED 331474
REFERENCE 20 (bases 39062 to 39170)
AUTHORS Denniston-Thompson, K., Moore, D.D., Kruger, K.E., Furlth, M.E. and Blattner, F.R.
TITLE Physical structure of the replication origin of bacteriophage lambda
JOURNAL Science. 198 (4321), 1051-1056 (1977)
MEDLINE 78054731
PUBMED 929187
REFERENCE 21 (bases 44467 to 44807)
AUTHORS Sklar, J.L.
TITLE Structure and function of two regions of DNA controlling the synthesis of prokaryotic RNAs
JOURNAL Thesis (1977)
JOURNAL 22 (sites)
REFERENCE 23 (bases 13 to 72; 48391 to 48502)
AUTHORS Adhya, S. and Gottesman, M.
TITLE Control of transcription termination
JOURNAL Annual review of biochemistry. 47, 967-996 (1978)
MEDLINE 78234064
PUBMED 354508
REFERENCE 24 (bases 13 to 72; 48391 to 48502)
AUTHORS Nichols, B.P. and Donelson, J.E.
TITLE 178-Nucleotide sequence surrounding the *cos* site of bacteriophage lambda DNA
JOURNAL Journal of virology. 26 (2), 429-434 (1978)
MEDLINE 78197067
PUBMED 666898
REFERENCE 24 (bases 37938 to 38016; 35589 to 35666)
AUTHORS Flashman, S.M.

Query Match	Best Local Similarity	Matches	98.8% Pred. No. 4.4e-283	DB 7	Length 48502	Gaps	Indels
1	aaagcgagaggtgtgtagcgagactctcgccaccgcgttcacgaagatcatgtgta	45005					
45064	AAACGGCAGAGGTTTAAACGCGACTCTCCACCCGCTTCACAGAGGACATGTA	45005					
61	aaagcgagaggtgtgtagcgagactctcgccaccgcgttcacgaagatcatgtgta	44945					
45004	AAAGGCGGAGGTAACATTAATTAATGAATTAACAGACGACAGTGGCTACAGT	180					
121	tgggtgtcgtctgctgagcgagagatgagacgctgtaacgttgcgttcgtcgc	44885					
44944	TGGGTTGTGCTGTGTGCGGCGGAGTACGCTTCACCATTTGGTGATCCGTTCT	240					
181	ttccggtatcgttataatcagacacaagaaagagactggtctaacaggtctgcgc	44825					
44884	TTCCGGRTTTCCTTAATTCACACACAGGAAGACAGTGGCTTAACAGGCTGCCG	300					
241	tcttcagatataagactcaatgactccttactgtgtgagatataaaatccogaac	44765					
44824	TCCTTCACGATTAATGACTCAATGCTTTACCTGTTGTGCGATTAATAAATCCGA	360					
301	cgctatgcagcgtctcaactatcctcggaaactgttllcggatgtcaatttcaga	44705					
44764	CGTTATGACGCGCTTMACTATTACTTCGCAACTGTTTGGGATTTGCTTTTGCAC	420					
361	tctgcgcgcaatgattgtagtcttcagagaatcgttcgtagagatgacaaactag	44645					
44704	TCGCGTCGCATGTTGAGTTCACAGACGATAGCTTGAGATGACCACTAGGGGAA	480					
421	gtagtaagcgcgcgtcttcatctcaactacacaaagagagaaatlaaccatcg	44585					
44644	GTAGTAACGCGCGCTCTTTTCAATGCTACCTCCACACGAGCAATTAACCATG	540					
481	gtcaattaccacaaatttcaataagtaataataatgacgttaataatgtgcac						

Db	44584	GTCAAAATTACCCAAATTTTATTCATNAATGCAATATCGCGCTTAATATGTTGGCAATCC	44525
QY	541	gtggcaatcatgctctgtaagctgtgaccgcattgaaatgtgtctgcatgtgacatcttc	600
Db	44524	GTCGCATATATGCTCTTAACCTGTGACCGCATTCAAAATGTTGCTCGATTGACATCTTC	44465
QY	601	ctttgtgcatctgcaaccacagagctcatatacagcgtgcttaaaagtcgtgccaagcttgg	660
Db	44464	TTTGTGCATTTGCAACCACGACGCTCATTAAGCGGCTTAAAGATGCGTACAGCGAGTGGG	44405
QY	661	cttgggtgaagcttggagattagcatctgtaaagcgcgatgctgctgtctgtgcatcct	720
Db	44404	TTGGGTAAAGTTTGGCATTTAGATGTACAGTGTACAGCGCGATGATGCTTGCGCTTGCGCATCCT	44345
QY	721	tgaatagcgcagagccttgcattcttcctgcgaactcttcctgaaacactctccccaagctc	780
Db	44344	TGAAATAGCGAGGCGCTTTGCATCTCCGACACTTTTTCGACAAACTCTCCCCACAGCTC	44285
QY	781	tgctttggcaataataaccgcaagggccttaccatggaactcttgatattgtgcccccg	840
Db	44284	TGTTTGGCAATATCAACCGACGCGCTGTAGCATGTGCATCTTGCAATCTTGCCCGCG	44225
QY	841	cgctgcgagcctagcgaaataatccgataaggaatgtgtgagcaacttgcagtaactt	900
Db	44224	CGTCGCGCACTACGGCAATATCGGCTATTAAGCGATGTGGAGACACTTCCAGTACCTT	44165
QY	901	tgactagtaattctctcaagctgccc	928
Db	44164	TGCCTTAGATTTCTTCACAGCTTTGCC	44137

FEATURES	source	BASE COUNT	ORIGIN	COMMENT
02453/c	9170 bp	DNA	SYN	29-MAR-1996
OCUS	Cloning vector lambda EMBL3, right arm.			
DEFINITION	002453			
ACCESSION	002453			
VERSION	002453.1	GI:413819		
KEYWORDS	Cloning vector lambda EMBL3.			
SOURCE	Cloning vector lambda EMBL3			
ORGANISM	artificial sequence: vectors.			
REFERENCE	1 (bases 1 to 9170)			
AUTHORS	Kitts,P.A.			
TITLE	CLONTECH Vectors On Disc version 1.3			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 9170)			
AUTHORS	Frischaut,A.M., Lehrach,H., Poustka,A. and Murray,N.			
TITLE	lambda replacement vectors carrying polynuker sequences			
JOURNAL	J. Mol. Biol. 170 (4), 827-842 (1983)			
MEDLINE	84064856			
REFERENCE	3 (bases 1 to 9170)			
AUTHORS	Kitts,P.A.			
TITLE	Direct Submision			
JOURNAL	Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA			
COMMENT	This sequence has been compiled from information in the sequence databases, published literature and other sources. If you suspect there is an error in this sequence, please contact CLONTECH's technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.			
FEATURES	Location/Qualifiers			
source	1. 9170			
	/organism="Cloning vector lambda EMBL3"			
	/db_xref="taxon:31783" 2389 t			
BASE COUNT	2534 a	1996 c	2251 g	
ORIGIN				
Query Match	98.7%	Score 921.6:	DB 12:	Length 9170:
Best Local Similarity	99.6%	Pred. No. 1,4e+282:		
Matches	924:	Conservative	Mismatches 4:	Indels 0: Gaps 0:


```

OY 1 aaacgcagaggtgtgttagcgcgaactctctgcacccgccttcacgaagtcagtgtla 60
Db 5732 AAACGCGAGAGGTTGTTAGCGGACCTCCGCCACCCGCTTTCAGGAAGGTCAGTGTTA
OY 61 aaagccgcagcgttaactaactaactaactaactaactaactaactaactaacta 5673
Db 5672 AAAGCGCGAGCGTAACTAATTAATGAGTTCAGGACAGACAGGCGCTACGGCTCAGTT
OY 121 tgggttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5613
Db 5612 TGGGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
OY 181 ttcggaatcgccttaactaactaactaactaactaactaactaactaactaacta 5553
Db 5552 TTCGGATTCGCTTAATTCAGACACAGGAAGAGACGCTGCTAACAGGCTGCGCAC
OY 241 tcttcacgattatcgccttaactaactaactaactaactaactaactaactaacta 5493
Db 5492 TCTTCACGATTATCGACTCAATAGCTTACTGTTGTCAGATATTAATAATCCGGAAC
OY 301 cgtatgcagagctcactaactaactaactaactaactaactaactaactaacta 5433
Db 5432 CGTTATGACAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
OY 361 tctgcctgcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5373
Db 5372 TCTGCTGCGATGTTGTGAGTTCAGACAGATGCTGCAAGTACCACTAGGCGGATCG
OY 421 gtaagaagccgcctctctctctctctctctctctctctctctctctctctctct 5313
Db 5312 GTACTAGCGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
OY 481 gtaaatatcccaattatcaatgaatgaatgaatgaatgaatgaatgaatgaatga 5253
Db 5252 GTCAAAATTAACCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
OY 541 gttgcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5193
Db 5192 GTGGCAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
OY 601 ttgtggaattgcacacacacacacacacacacacacacacacacacacacacacac 5133
Db 5132 TTGTGCGATTTGACACACACACACACACACACACACACACACACACACACACACAC
OY 661 ttgggttaagtttgggttaagtttgggttaagtttgggttaagtttgggttaagtt 5073
Db 5072 TTGGGTTAAGTTTGGGTTAAGTTTGGGTTAAGTTTGGGTTAAGTTTGGGTTAAGTTT
OY 721 tgaatagccagcgccttcgacatcttcgacatcttcgacatcttcgacatcttcgac 5013
Db 5012 TGAATAGCCGAGCGCTTTCGACTCTTCTGCACTCTTCTGCACTCTTCTGCACTCTT
OY 781 tgtttggaataataacacacacacacacacacacacacacacacacacacacacac 4953
Db 4952 TGTTTGGCAATATACACACACACACACACACACACACACACACACACACACACAC
OY 841 cgtgcgcacactaagcgaataacacacacacacacacacacacacacacacacacac 4893
Db 4892 CGTGCACGCACTAAGCAATATCCGCAATATCCGCAATATCCGCAATATCCGCAATAT
OY 901 tgccttagtatttcttcaagtgcgc 928
Db 4832 TGCCCTAGTATTCTTCCTCAAGCTTTGCC 4805

```

```

RESULT 9
LOCUS XXU02427/c 9205 bp DNA
DEFINITION Cloning vector lambda EMBL3 SP6/77, right arm.
ACCESSION U02427
VERSION 002427.1 GI:413793
KEYWORDS Cloning vector lambda EMBL3 SP6/77.
SOURCE

```

```

ORGANISM Cloning vector lambda EMBL3 SP6/77
REFERENCE Artificial sequence; vectors.
AUTHORS Kites, P.A.
TITLE Clontech Vectors on Disc version 1.3
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 9205)
TITLE Kites, P.A.
JOURNAL Direct Submission
COMMENT 1020 East Meadow Circle, Palo Alto, CA 94303, USA
This sequence has been compiled from information in the sequence
databases, published literature and other sources. If you suspect
there is an error in this sequence, please contact CLONTech's
Technical Service Department at (415) 424-8222 or (800) 662-2566,
extension 3 or E-mail TECH@CLONTech.COM.
FEATURES
Source
Location/Qualifiers
1..9205
/organism="Cloning vector lambda EMBL3 SP6/77"
/db_xref="taxon:31784"
BASE COUNT 2546 a 2002 c 2253 g 2404 t
ORIGIN
Query Match 98.7%; Score 921.6; DB 12; Length 9205;
Best local Similarity 99.6%; Pred. No. 1.4e-282;
Matches 924; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 1 aaacgcagaggtgtgttagcgcgaactctctgcacccgccttcacgaagtcagtgtla 60
Db 5767 AAACGCGAGAGGTTGTTAGCGGACCTCCGCCACCCGCTTTCAGGAAGGTCAGTGTTA
OY 61 aaagccgcagcgttaactaactaactaactaactaactaactaactaactaacta 5708
Db 5707 AAAGCGCGAGCGTAACTAATTAATGAGTTCAGGACAGACAGGCGCTACGGCTCAGTT
OY 121 tgggttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5648
Db 5647 TGGGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
OY 181 ttcggaatcgccttaactaactaactaactaactaactaactaactaactaacta 5588
Db 5587 TTCGGATTCGCTTAATTCAGACACAGGAAGAGACGCTGCTAACAGGCTGCGCAC
OY 241 tcttcacgattatcgccttaactaactaactaactaactaactaactaactaactaacta 5528
Db 5527 TCTTCACGATTATCGACTCAATAGCTTACTGTTGTCAGATATTAATAATCCGGAAC
OY 301 cgtatgcagagctcactaactaactaactaactaactaactaactaactaacta 5468
Db 5467 CGTTATGACAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
OY 361 tctgcctgcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5408
Db 5407 TCTGCTGCGATGTTGTGAGTTCAGACAGATGCTGCAAGTACCACTAGGCGGATCG
OY 421 gtaagaagccgcctctctctctctctctctctctctctctctctctctctctct 5348
Db 5347 GTAGTAAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
OY 481 gtaaatatcccaattatcaatgaatgaatgaatgaatgaatgaatgaatgaatga 5288
Db 5287 GTCAAAATTAACCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
OY 541 gttgcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5228
Db 5227 GTGGCAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
OY 601 ttgtggaattgcacacacacacacacacacacacacacacacacacacacacacac 5168
Db 5167 TTGTGCGATTTGACACACACACACACACACACACACACACACACACACACACACAC
OY 661 ttgggttaagtttgggttaagtttgggttaagtttgggttaagtttgggttaagtt 720

```

```

Db 5107 TTGGGTAAGTTTGGATFAGCATCTCACAGCGGATATGCTGGCTTGGCATCT 5048
OY 721 tgaatagcgaagccttgcattcttcgcacatcttcgcacacatctcccccagctc 780
Db 5047 TGAATAGCCGACGCTTGGATCTCCGACACTCTTCTGACAACTTCGCCACAGCTC 4988
OY 781 tgtttgcaatcatcaacgcgaagcctgtacacatgcaatctctgcacatcttgcacccg 840
4987 TGTTTGGCAATFATCAACCGACGCGCTGTACATGCAATCTCTGCACTTGGCCCGG 4928
OY 841 cgtcgagcgaactgagcgaatcatccgataagcgaatgttgcagacacttgcacatctt 900
Db 4927 CGTCGGGACATACGCGCATATATCCGATTAAGCAATGTTGGCAGCACTTGACCTT 4868
OY 901 tgccttagtattccttcaagctgccc 928
Db 4867 TGCTTAGTATTTCCCTCAAGCTTGGC 4840

```

RESULT 10
LAMILPR/c 735 bp DNA PHG 28-APR-1993

LOCUS Bacteriophage lambda late transcription promoter region DNA.
DEFINITION M38294
ACCESSION M38294
VERSION M38294.1 GI:215179

KEYWORDS Bacteriophage lambda DNA.
SOURCE Bacteriophage lambda
ORGANISM viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;

REFERENCE 1 (bases 1 to 735)
Petrov, N.A., Karginov, V.A., Mikryukov, N.N., Serpinsky, O.I.,
AUTHORS Kravchenko, V.V. and Vasilenko, S.K.

TITLE Nucleotide sequence of a bacteriophage lambda DNA fragment
containing the late transcription promoter P-R

JOURNAL Bioorg. Khim. 5, 1873-1875 (1979)
LOCATION/Qualifiers

FEATURES 1..735
source /organism="Bacteriophage lambda"
/db_xref="taxon:10710"

BASE COUNT 202 a 173 c 192 g 168 t
ORIGIN

Query Match 70.4%; Score 657.4; DB 7; Length 735;
Best Local Similarity 98.2%; Pred. NO. 2.9e-198; Indels 7; Gaps 6;
Matches 728; Conservative 0; Mismatches 6;

```

OY 66 ccgacgagtaactactactaatgaattcagacagacagctgagctacgagctgaatttggt 125
Db 735 CCGACGAGCTACTACTACTAATGAATTTCAGACAGACAGTGGCTACGGCT AGTTTGGGT 677
OY 126 tgtccttctgctggcgagcgaatgagcctgtacgattgtgtatccgcttctcgcg 185
Db 676 TGTCTGCTGCTGGCGGCGCATGACGCTGTACGCACTTTGGTGATCCGGTTCTGCTCCG 617
OY 186 gtattcgttaattcagacacagcgaagagcactgtgtacacagcctgcgcgactctc 245
Db 616 GTATTGCTTAACTTACGACACAGCAAGAGACTGGCTAACCAAGCTGCGGACTCTTC 557
OY 246 acgattatcgaaccgaatgtcttcaactcgtgtgacagataaataatccgaacacgta 305
Db 556 ACGATTATCGACTCAATCTCTTACCTGTGTGACAGATATAAATAATCCGAAACCGTAA 497
OY 306 tgcagagcttaactactactcgcgaactgttgcgagatgtgacatttgcagacactctgc 365
Db 496 TGCAGGCTTAACTACTACCTGCGCAACTGTTCCGGATTCGATTTTGCACACTCTCTCCG 437
OY 366 ctgcagatgtctgagatcagaacgaatgtagaagtaaccaactagcaggaatcgtgagt 425
Db 436 CTGCGATGCTTGGAGTTCCAGACGATACGCTGAAGTGAACCAACTAGGCGGAATCGTAGT 377

```

```

OY 426 aagcgccgctcttctcatctcactacacacagcagcgaatgaaccacatcgttgatcaa 485
Db 376 AAGCGCCGCTCTTTCATCTCACTACACACAGAGGAAATTAACCATCTTGAGTCAA 317
OY 486 attaccaattatctcaataagtaataatcagcctgaatattgtgcacatcgtgc 545
Db 316 A-TTACCCCAATTTATTAATATGCA--ATCATGCCGTTAATATG-TGCCATCCCTGGC 261
OY 546 aatcatcgtcgaactgtagcgcgacatcaaatgtgtctgcagattgactcttctgt 605
Db 260 GATCATGCTGCTAAGCTGATGATGATCAAAATGTGTCTGCATTTGACTCTCTTTGG 201
OY 606 ggcattgcacaccagagcgtatatacagcgcgttaacagtgctgacacagtggttg 665
Db 200 GGCATTGCACACCAAGCGCTATACAGCGGCTTAAAGTCGTAACAGAGCTGGGTTGG 141
OY 666 taagtttgagtaagcagcgtacagcagcagcagcagcagcagcagcagcagcagc 725
Db 140 TAAAGTTTGGGATTTAGCATTCACAGCCGATA-GTTCGCTTGTGTCATCTTGAAT 82
OY 726 agccagcgccttgcacatcttcgcacactcttcgcacactctcccccagcctctt 785
Db 81 AGCCGAGCGCTTTCATCTTCGCACTCTTCTCGACAACTCCGCCACAGCTCTGTT 22
OY 786 tgg--caatcaacgcagcag 805
Db 21 TGGACATATTCAGCGCGCGG 1

```

RESULT 11
LAMILPR/c 751 bp DNA PHG 28-APR-1993

LOCUS Bacteriophage lambda Q protein gene, complete cds.
DEFINITION M38285
ACCESSION M38285
VERSION M38285.1 GI:215194

KEYWORDS Q protein.
SOURCE Bacteriophage lambda DNA.
ORGANISM Bacteriophage lambda
viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;

REFERENCE 1 (bases 1 to 751)
Petrov, N.A., Serpinsky, O.I., Mikryukov, N.N., Karginov, V.A. and
AUTHORS Kravchenko, V.V.

TITLE Nucleotide sequence of the bacteriophage lambda DNA region
containing gene Q

JOURNAL Bioorg. Khim. 8, 561-563 (1982)
LOCATION/Qualifiers

FEATURES 1..751
source /organism="Bacteriophage lambda"
/db_xref="taxon:10710"

BASE COUNT 212 a 168 c 209 g 162 t
ORIGIN

Query Match 48.0%; Score 448.2; DB 7; Length 751;
Best Local Similarity 98.1%; Pred. NO. 1.7e-131; Indels 1; Gaps 1;
Matches 464; Conservative 0; Mismatches 8;

```

OY 456 aacgagcgaatcaaccatcgttgagtaaatcttaaccaattatcaataagtaata 515
Db 751 ACCGAGGAATTAACCATCTGTTGAGTCAAAATTTTACCAATTTTATCAATTAAGTCAATA 692
OY 516 tcagcgttaataatgttgcacatcgttgcaatcatcgtcgtcaagtgtagcgcacatca 575

```


50.. .535
/note="gp1"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase small subunit"
/protein_id="AA030353.1"
/db_xref="GI:6863112"
/translation="MADKIRSDSSAAVQAMKNAAYDTIIDPSSHGLEKKAEPFMD
NIRSKALDSWTPADLLAAVEELANNOIYTLRLKDRLEERIKSEPEDELLKDLKQI
VEIORTLIACRDLQIHSHATGESRQDKKRNQDRDARTNKNEHODDNLIAPKH
G"
542.. .2056
/note="gp2"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase large subunit"
/protein_id="AA030354.1"
/db_xref="GI:6863113"
/translation="MTRGEVIAIERFPCIVPEGKLIQGPMLRDTPOKEFTIAYVDP
AGDMALILEIARKNGKTLIGTILAHVGEPAVONTQIVSALSREQAIVPRLAYK
MVNINRLEIYHITPESGKLIIGLPCVVEKALSAGKTTGLSPILALIDETQVNG
PODDFDATITAGAHENPLIVISTQANADALLSIMIDDAKSKDPHIVCHYTEP
KODDISRESMLANPALGTFRSEKDMARQAEKAGMPSEPTFRNLINORVSTVS
FISSWELCGEMPINTPRKWAYGJDSARNDALTYIAGEADGVWVFPEFWPQK
TLEERTIDPAIVMREGLIRTPGASVYSVAVADIAETIGDGLTSLAFDMRL
DOFRKDDALIGSLPLVERGGFKDMGPVDTLSLMLNGRVRRGHHPVLTMCANAV
VYKDAAGRKLDKSKATGRIDGMVMTWSVGAANGVTEGGDFDFIRPLISM"
2056.. .3330
/note="gp3"
/codon_start=1
/transl_table=11
/product="head portal protein"
/protein_id="AA030355.1"
/db_xref="GI:6863114"
/translation="MEEPKYITIDLRITNGMNRALQSWFVGGHLYVINGSGTGPVSAH
GHIGDSINDERLIQISTVRCVSLISTVACPLIDVFPTDONDNRKKVVDLSNPLAL
LRSPMOTVADPEFAMTMOICFYGNALVLYRNSAGVLSILPLGSANMDYKLVGK
KVYRFORSEVADFSOKEIFHLKGFETGLVSLTAPACKSAGVAAAMDQORDF
ANEMKSSQLISTGEKVIITEOOSVVEENFEIAGGEPVKRLNITILEAGSTSAIGTPO
DAEMKARFOVSELRAPRGVPLVGVGEKSTWSGILBOONLGLIOTYLOPYSIM
ENSIGRWLIPAKDVRIHAENHLDGLRGDSASRAAFKAMGAGLRTINERTRDNL
PPLGGGVAMROSOYVPIIDTCTNKEPNNCA"
3348.. .4025
/note="gp4"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="head maturation protease"
/protein_id="AA030356.1"
/db_xref="GI:6863115"
/translation="MPETVITISPDETEIKFTGDKQGIIFBYASVFNNTDSQDILL
PGAFKALANOTRKVAFENRIMELPVGMKDSIADEKGLVLRQOUTPRHSGAIAMKS
AAHQGTVEGMSVGFVSXKDDYTIIPGRIFKNIOALREI SVCTEPAPBOAGIAAMKS
VDGIETIRDEVNMLRDSVGLIKSQAVGLIARFSAIRSESEGDGNEAINALLOSIS
FPSNLCK"
4028.. .5185
/note="gp5"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="major capsid subunit precursor"
/protein_id="AA030357.1"
/db_xref="GI:6863116"
/translation="MSRLAIKKAIESQOKMTQLFPAOKAIEISGQVSKOLOSIM
KQOEILTSKGRILDEOKLAISGANGEEKSFSERAAEELIKSDGKOGTFRACKFN
KISGSDASAGSLIOPMOTFGLIMPGLRILIRDLAAGRTSSNALTIVBEPTFNA
DYVAEKALIPESDITFSKQVANKRTIAHWVQASROYMDAPMLQSYINNNLIGLALX
EEGOLINGGTGDNLEGLNKAIVATIDISINATNGIRADIIAHAYOYTESGFSSGIV
LNRDHNHIALTKDEGRYIFGGPOATISIMGGLPVYTRKQAAGAYTEVGGFDMASQ
VMDRMATVEYSEDRDNFVKNMLTILCEERLAIHRYPTALIKGTSSGS"
5219.. .5545

50.. .535
/note="gp1"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase small subunit"
/protein_id="AA030353.1"
/db_xref="GI:6863112"
/translation="MADKIRSDSSAAVQAMKNAAYDTIIDPSSHGLEKKAEPFMD
NIRSKALDSWTPADLLAAVEELANNOIYTLRLKDRLEERIKSEPEDELLKDLKQI
VEIORTLIACRDLQIHSHATGESRQDKKRNQDRDARTNKNEHODDNLIAPKH
G"
542.. .2056
/note="gp2"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase large subunit"
/protein_id="AA030354.1"
/db_xref="GI:6863113"
/translation="MTRGEVIAIERFPCIVPEGKLIQGPMLRDTPOKEFTIAYVDP
AGDMALILEIARKNGKTLIGTILAHVGEPAVONTQIVSALSREQAIVPRLAYK
MVNINRLEIYHITPESGKLIIGLPCVVEKALSAGKTTGLSPILALIDETQVNG
PODDFDATITAGAHENPLIVISTQANADALLSIMIDDAKSKDPHIVCHYTEP
KODDISRESMLANPALGTFRSEKDMARQAEKAGMPSEPTFRNLINORVSTVS
FISSWELCGEMPINTPRKWAYGJDSARNDALTYIAGEADGVWVFPEFWPQK
TLEERTIDPAIVMREGLIRTPGASVYSVAVADIAETIGDGLTSLAFDMRL
DOFRKDDALIGSLPLVERGGFKDMGPVDTLSLMLNGRVRRGHHPVLTMCANAV
VYKDAAGRKLDKSKATGRIDGMVMTWSVGAANGVTEGGDFDFIRPLISM"
2056.. .3330
/note="gp3"
/codon_start=1
/transl_table=11
/product="head portal protein"
/protein_id="AA030355.1"
/db_xref="GI:6863114"
/translation="MEEPKYITIDLRITNGMNRALQSWFVGGHLYVINGSGTGPVSAH
GHIGDSINDERLIQISTVRCVSLISTVACPLIDVFPTDONDNRKKVVDLSNPLAL
LRSPMOTVADPEFAMTMOICFYGNALVLYRNSAGVLSILPLGSANMDYKLVGK
KVYRFORSEVADFSOKEIFHLKGFETGLVSLTAPACKSAGVAAAMDQORDF
ANEMKSSQLISTGEKVIITEOOSVVEENFEIAGGEPVKRLNITILEAGSTSAIGTPO
DAEMKARFOVSELRAPRGVPLVGVGEKSTWSGILBOONLGLIOTYLOPYSIM
ENSIGRWLIPAKDVRIHAENHLDGLRGDSASRAAFKAMGAGLRTINERTRDNL
PPLGGGVAMROSOYVPIIDTCTNKEPNNCA"
3348.. .4025
/note="gp4"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="head maturation protease"
/protein_id="AA030356.1"
/db_xref="GI:6863115"
/translation="MPETVITISPDETEIKFTGDKQGIIFBYASVFNNTDSQDILL
PGAFKALANOTRKVAFENRIMELPVGMKDSIADEKGLVLRQOUTPRHSGAIAMKS
AAHQGTVEGMSVGFVSXKDDYTIIPGRIFKNIOALREI SVCTEPAPBOAGIAAMKS
VDGIETIRDEVNMLRDSVGLIKSQAVGLIARFSAIRSESEGDGNEAINALLOSIS
FPSNLCK"
4028.. .5185
/note="gp5"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="major capsid subunit precursor"
/protein_id="AA030357.1"
/db_xref="GI:6863116"
/translation="MSRLAIKKAIESQOKMTQLFPAOKAIEISGQVSKOLOSIM
KQOEILTSKGRILDEOKLAISGANGEEKSFSERAAEELIKSDGKOGTFRACKFN
KISGSDASAGSLIOPMOTFGLIMPGLRILIRDLAAGRTSSNALTIVBEPTFNA
DYVAEKALIPESDITFSKQVANKRTIAHWVQASROYMDAPMLQSYINNNLIGLALX
EEGOLINGGTGDNLEGLNKAIVATIDISINATNGIRADIIAHAYOYTESGFSSGIV
LNRDHNHIALTKDEGRYIFGGPOATISIMGGLPVYTRKQAAGAYTEVGGFDMASQ
VMDRMATVEYSEDRDNFVKNMLTILCEERLAIHRYPTALIKGTSSGS"
5219.. .5545

50.. .535
/note="gp1"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase small subunit"
/protein_id="AA030353.1"
/db_xref="GI:6863112"
/translation="MADKIRSDSSAAVQAMKNAAYDTIIDPSSHGLEKKAEPFMD
NIRSKALDSWTPADLLAAVEELANNOIYTLRLKDRLEERIKSEPEDELLKDLKQI
VEIORTLIACRDLQIHSHATGESRQDKKRNQDRDARTNKNEHODDNLIAPKH
G"
542.. .2056
/note="gp2"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase large subunit"
/protein_id="AA030354.1"
/db_xref="GI:6863113"
/translation="MTRGEVIAIERFPCIVPEGKLIQGPMLRDTPOKEFTIAYVDP
AGDMALILEIARKNGKTLIGTILAHVGEPAVONTQIVSALSREQAIVPRLAYK
MVNINRLEIYHITPESGKLIIGLPCVVEKALSAGKTTGLSPILALIDETQVNG
PODDFDATITAGAHENPLIVISTQANADALLSIMIDDAKSKDPHIVCHYTEP
KODDISRESMLANPALGTFRSEKDMARQAEKAGMPSEPTFRNLINORVSTVS
FISSWELCGEMPINTPRKWAYGJDSARNDALTYIAGEADGVWVFPEFWPQK
TLEERTIDPAIVMREGLIRTPGASVYSVAVADIAETIGDGLTSLAFDMRL
DOFRKDDALIGSLPLVERGGFKDMGPVDTLSLMLNGRVRRGHHPVLTMCANAV
VYKDAAGRKLDKSKATGRIDGMVMTWSVGAANGVTEGGDFDFIRPLISM"
2056.. .3330
/note="gp3"
/codon_start=1
/transl_table=11
/product="head portal protein"
/protein_id="AA030355.1"
/db_xref="GI:6863114"
/translation="MEEPKYITIDLRITNGMNRALQSWFVGGHLYVINGSGTGPVSAH
GHIGDSINDERLIQISTVRCVSLISTVACPLIDVFPTDONDNRKKVVDLSNPLAL
LRSPMOTVADPEFAMTMOICFYGNALVLYRNSAGVLSILPLGSANMDYKLVGK
KVYRFORSEVADFSOKEIFHLKGFETGLVSLTAPACKSAGVAAAMDQORDF
ANEMKSSQLISTGEKVIITEOOSVVEENFEIAGGEPVKRLNITILEAGSTSAIGTPO
DAEMKARFOVSELRAPRGVPLVGVGEKSTWSGILBOONLGLIOTYLOPYSIM
ENSIGRWLIPAKDVRIHAENHLDGLRGDSASRAAFKAMGAGLRTINERTRDNL
PPLGGGVAMROSOYVPIIDTCTNKEPNNCA"
3348.. .4025
/note="gp4"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="head maturation protease"
/protein_id="AA030356.1"
/db_xref="GI:6863115"
/translation="MPETVITISPDETEIKFTGDKQGIIFBYASVFNNTDSQDILL
PGAFKALANOTRKVAFENRIMELPVGMKDSIADEKGLVLRQOUTPRHSGAIAMKS
AAHQGTVEGMSVGFVSXKDDYTIIPGRIFKNIOALREI SVCTEPAPBOAGIAAMKS
VDGIETIRDEVNMLRDSVGLIKSQAVGLIARFSAIRSESEGDGNEAINALLOSIS
FPSNLCK"
4028.. .5185
/note="gp5"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="major capsid subunit precursor"
/protein_id="AA030357.1"
/db_xref="GI:6863116"
/translation="MSRLAIKKAIESQOKMTQLFPAOKAIEISGQVSKOLOSIM
KQOEILTSKGRILDEOKLAISGANGEEKSFSERAAEELIKSDGKOG

Query Match	46.7%	Score 436.2;	DB 7;	Length 40751,
Best Local Similarity	95.1%;			
Matches 450;	Conservative			

		Indels	Gaps
QY	456	aagcagcgaatlaaacccatcgttgaagccaaatltaaccaatlttlaataaagicaata	515
Db	36934	AACGAGTCGATTTTTCATGCTGGTGGTGAATTAATTAACCACTTTATTTCAAAAAAGTCATA	36872
QY	516	taicgcgtlaatatgtlttgcatccgttgcaataicgcgtacagctylgacgcgcatca	575
Db	36874	TCATGCCGTTAAATATGTTGCGATCCGGTGCAATATGCTGCTAAAGTCCGACCCGATTC	36815
QY	576	aaatgtgtctggaattgaactcttcttlttggcgaattgaacacacgcagacgctcaataagcg	635
Db	36814	AAATGTTGTCGTGGATGATGACTTCTCTTGGCATTTGACACCGACGAGCCGATTAACGG	36755
QY	636	gcttaacagltgcgtaccagatgggttgaglaagttltggatlagatcgttcaagcgc	695
Db	36754	GCTTAAACAGTCGCTCACCAGATGGGTGGTGAAGGTTTGGATTAAGCATCTGTAAAGCGC	36695
QY	696	gatatctgcgcgttgcgtgcgaactccatlgaaatagcagacgcttgaactcttcgcgaactct	755
Db	36694	GATATGCGGCGCTTGCTGCGGATCGATCGAATACCGACGCGCTTTGCATCTTCCGCACTTT	36635
QY	756	tctgcgaacacttcccacaagctcgttlttggcaatataaacgcgacgcgtctaacat	815
Db	36634	TCTGCACACACTCTCCCACTGCTCTGTTTTGGCAATATCAACGCGCCGCGCACTTACGTT	36575
QY	816	ggcaactctgcaatcttgcctccgcgcgttcgcgcgaactcgaagaaataicgcgtaaagca	875
Db	36574	GGCATATTCGATCTTGGCGCCCGGTGCTCCCGGCACTAAGGCAATTAATTCGCAATTAAGCA	36515
QY	876	atgttgcggaacttgcagatcacttgccttaglaatttcccttcaagctgcgcg	928
Db	36514	ATGTTTCGGAACCTTGCAGTACCTTTGCTTGATTTTCCCTTCAAGCTTTGCC	36462

RESULT	14				
LOCUS	AF335538/c				
DEFINITION	AF335538	38297 bp	DNA	circular	
ACCESSION	Bacteriophage HK620,	complete genome.		PHG	03-APR-2001
VERSION	AF335538				
KEYWORDS	AF335538.1	GI:13517559			
SOURCE	Bacteriophage HK620.				
ORGANISM	Bacteriophage HK620				
	viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae				

REFERENCE AUTHORS TITLE	JOURNAL REFERENCE AUTHORS TITLE	FEATURES SOURCE
(bases 1 to 38297) Clark,A.J., Inwood,W.B., Cloutlier,T. and Dhillon,T.S. Nucleotide Sequence of Coliphage HK620 and the Evolution of Lambdaoid Phages	J Mol. Biol. (2001) In press 2 (bases 1 to 38297) Clark,A.J., Inwood,W.B., Cloutlier,T. and Dhillon,T.S. Direct Submission	Submitted (11-JAN-2001) Life Sciences Division, Lawrence Berkeley Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA Location/Qualifiers 1. 38297
		/organism="Bacteriophage HK620" /db_xref="taxon:155148" /lab_host="Escherichia coli H, strain 2158" complement(join(1..23,38289..38297)) /note="terminator for prophage excision transcript: XIS" /gene="int" /note="pkA" complement(50..1207)
terminator		
gene		
CDS		

```

/feature="site-specific recombination"
/codon_start=1
/transl_table=11
/product="Integrase"
/protein_id="AAK28849.1"
/db_xref="GI:13517560"
/translation="MLTVKQIEAARKREKPYLLDSNGLIYLPVSKKVMQIRKYL
GKEKILGYKPYPLMTLOEARDKMAARKDI SVGIDIPYAKKAPENNNSSPAITKEVY
HKQVWSSGYGTETELAKMFDDDLKIPAPADLADAMKGYRKKNPPLDPOIIFAKALATSGSK
RRCCEVRRVIAIVYGRAKTINPAPADLADAMKGYRKKNPPLDPOIIFAKALATSGSK
IYSLIAKTIITRLTLRKTELRSQMKMNVDFENRIITIDAVMGKRIHVAPSDQVYV
LTLTSLSTYKPVSEFVFPARNDKKPICENNVLIVTQITDGLSEGHGFRFEESTIIM
NEHMPADALEVQLAHANGSVRGIIYHNAQYLDKRRRMMQADMWIDKEVE"
829..839
/note="atp"
1207..1232
/note="atp"
complement(1224..1273)
/note="prophage excision"
1232..1253
/note="atp"
complement(1324..1464)
/gene="hxab"
complement(1324..1464)
/gene="hxab"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAK28850.1"
/db_xref="GI:13517561"
/translation="MKTLGKIFLTLSKKNHNSQSPASSNCLQGTPIISSLSQSPVPR
IP"
1366..1383
/note="atp core"
1441..1463
/note="atp"
complement(1457..1657)
/gene="hxac"
complement(1457..1657)
/gene="hxac"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK28851.1"
/db_xref="GI:13517562"
/translation="MOHELOPDSLVDLKFIMADGFGKFTFYDIRIKSDLPKAVIHG
BARMLYDHOEFKNKILSLANG"
complement(1664..1702)
/function="phage integration"
complement(1689..1701)
/bound_molecy="cII protein"

```

```

gene      /function="Promoter activation to establish lysogeny"
CDS       complement(1789..1968)
           /gene="hkad"
           /complement(1789..1968)
           /note="similar to eag of Bacteriophage P22"
           /codon_start=1
           /transl_table=11
           /product="unknown"
           /protein_id="AAK28852.1"
           /db_xref="GI:13517563"
           /translation="MSCPKGSGNIAKEKTMGWSGDYVCCDCGTNSKAFGRGKN
           EVKINSESGNEKS"
gene      complement(2065..2598)
           /gene="hkag"
           /complement(2065..2598)
           /codon_start=1
           /transl_table=11
           /product="unknown"
           /protein_id="AAK28853.1"
           /db_xref="GI:13517564"
           /translation="MTITKRTIFLTKNPLNDGILTRGEMEVARIALSLGAEPSVQ
           TYNLELIEGMEVSDVSTCDADGNKRGCTVPEALFLDPAKNGYILLVQDAPENDV
           NGSFVTPDGMVISCSDRMPDYKTAVALAVEPDRKGRMKMATYIPGHPDANDGWIIIP
           GASWPKSHWMLPEPPL"
gene      complement(2601..2792)
           /gene="hkaf"
           /complement(2601..2792)
           /codon_start=1
           /transl_table=11
           /product="unknown"
           /protein_id="AAK28854.1"
           /db_xref="GI:13517565"
           /translation="MDGOISIVRPGACDREIRITRIILAMKAITALTIPENLALALT
           GKSPLPELKLBNVEIKV"
gene      complement(2794..3510)
           /gene="hkag"
           /complement(2794..3510)
           /note="similar to I0140 of Bacteriophage 93W and ead of
           Bacteriophage P22"
           /codon_start=1
           /transl_table=11
           /product="unknown"
           /protein_id="AAK28855.1"
           /db_xref="GI:13517566"
           /translation="MSKIDVQALREKAKETSGVSLSEYGEERFDADALLHREVVGY
           LPCRLEGAHPESGDEDEFOEQOANAEFLAANPATIVALLDELRRKOYIKRDOE
           NEIALIVGRLEVELEKDSKIANLTBERDALREGEGDARHSNTAAADIYFOLVEE
           CEIPAGGSIVYVDNREKLEAARIRIALEDRETLIPERSMLHTDHPDYQYVMA
           KYVSEYIAIRAGIRIKGE"
gene      complement(3507..3989)
           /gene="hkaf"
           /complement(3507..3989)
           /codon_start=1
           /transl_table=11
           /product="unknown"
           /protein_id="AAK28856.1"
           /db_xref="GI:13517567"
           /translation="MKMSLIEMDGLFKGKCPRLKYNETNAEYLVAKPELRSKLE
           TALBGSACITIDNLEAKCAKMAAENTSLKSEFNDPCEEFSEMEDDYTTPTT
           DAFLAEVRAQGVEMKADNLDNVADDADERGCFDYAAKFLRSEASSVRLPADOLRGGSSO"
gene      complement(3986..4150)
           /gene="hkai"
           /complement(3986..4150)
           /codon_start=1
           /transl_table=11
           /product="unknown"

```

```

terminator /protein_id="AAK28857.1"
            /db_xref="GI:13517568"
            /translation="MRGLAYNPGLIPAEKITIKQKVPMPSPRELLKRNPSFVYQNKY
            LNMWRSGKK"
gene      complement(4045..4057)
           /note="UL3"
           /complement(4161..4457)
           /gene="hkad"
           /complement(4161..4457)
           /gene="hkaj"
           /codon_start=1
           /transl_table=11
           /product="unknown"
           /protein_id="AAK28858.1"
           /db_xref="GI:13517569"
           /translation="MPAPLYGADDPRLKSGNSVSEVLDFKRNKYNDRINSLPOETKEEK
           EFRHCIMLAKEEREKRIYQISIRPKRATYTHPEYIDPLRMYHSRYAISND"
gene      complement(4481..5068)
           /gene="hkak"
           /complement(4481..5068)
           /codon_start=1
           /transl_table=11
           /product="unknown"
           /protein_id="AAK28859.1"
           /db_xref="GI:13517570"
           /translation="MSNVIITYNEBALTAQGGFINTGSAHITITTEAELOSEKGA
           KFIIEFGSESDGRIQYLSVQKNDGYNKRGANVYVYAMGCGAGIGLOIHOHVSASK
           FVAPEPHGKIGLVLOKVTNKKGTGADSYOMEIPIPIAOTGOTLEKKAQKOPETI
           ANMYASIKDKDNRSKNVSONHADYDYSONDYPPE"
gene      complement(5065..5745)
           /gene="hkai"
           /complement(5065..5745)
           /codon_start=1
           /transl_table=11
           /product="unknown"
           /protein_id="AAK28860.1"
           /db_xref="GI:13517571"
           /translation="MGTATLIGESGTGKSTSMRNINPEEAILIKRIGKPIPFKSKFW
           IAMDRAKRGVYTTDKWDVYVAVIKRAHEGRIYVDPQVMSNEMRSEKSEK
           DKFTEIGHHAWVIAAODAPDRIAYFLAHEETPMGRVYKMTIKGKIDETIVGM
           FTIVRLTLRDQFFFTKNGADTVSPMGMFSPNSIDNDLSFVATYCDYGGINN
           HOIKENAA"
gene      complement(5754..5942)
           /gene="hkam"

```

```

Query Match 46.0%; Score 429.8; DB 7; Length 38297;
Best Local Similarity 94.3%; Pred. No. 1.4e-125;
Matches 446; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```

```

QY 456 aacgagcgaattacccatcgctgagtcgaattacccaatcttttcaataagtcata 515
    |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 15617 AACGAGTCGATTTTCATCGTGGTCAATTTACCCACTTATATCAAAAAGTCATA 15558

QY 516 tcatgcgttaatatgttgcgcatcgcggaatcatgctgaacgctgacgcattca 575
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 15557 TCATGCCGTAAAGATGTCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 15498

QY 576 aatgtgtctgcgactgactctcttgcgcatcgcggaatcatgctgaacgctgacgc 635
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 15497 AATGTGTCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 15438

QY 636 gcttaacagtgctgacgagtgctgagtgctgagtgctgagtgctgagtgctgagtgct 695
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 15437 GCTTAACAGTGCCTGACCGAGTGCTGAGTGAAGTGTGGGATTAAGCATGCTTAAC 15378

QY 696 gatactgcgctgctgctgcatctctgaatagcgaagccttgcactcttcgactctt 755
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 15377 GATATGCTGCGCTTGGCATCTTGAATFACCGACGCTTTCGATCTTCGACGACTT 15318

QY 756 tctgcacaactcccccacagctctgtttgcaatatacaacgcgacgcttacctat 815
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE
1	Hendrix, R.W.	Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages	J. Mol. Biol.	20328598
2	Hendrix, R.W., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F., and	Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages	J. Mol. Biol.	20328598

PUBMED 10860721
 REFERENCE 2 (bases 1 to 39732)
 AUTHORS Jadhav, R., Ford, M. E.,
 TITLE Hendrich, R. W.
 JOURNAL Direct Submission
 Submitted (01-JUN-1998) Pittsburgh Biocenter
 of Biological Sciences, University of Pittsburgh, PA
 15260, USA

FEATURES	Location/Qualifiers
source	1. .39732

```
1...39732
/organism="Bacteriophage HK57"
/db_xref="taxon:37534"
50..535
/gene="1"
50..535
/gene="1"
/gene="1"
Note="Gnl"
```

```

/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase small subunit"
/db_xref="GI:6901594"
/translation="MADRRISDSSAAVAQMAAAVDITDPSPHAGLEKKAEPKWH
NLSKSLDSSWPAADLLAAVELAAVLAANNOLTVYLRKDIKREERISGEDESLIKDKLRK
VELQRTIIILAKRDLDIHSHATNGESRQDKRKNONDRDARNTKNEHQDQDDNLIAPPK
G"
542. .556
/gene="2"
542. .2056
/gene="2"
/note="gp2"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="terminase large subunit"
/protein_id="AAF31098.1"
/db_xref="GI:6901595"
/translation="MPRGSEVIAFLERFCTIVECKLIGQPKRLDFQKDFILAVYDNP
AGTMALISLARKKNGKGTGLIAGLIALHLVGPAPVANTOYSGALSRQDAIVPNLAVK
FQNDPILQETVETHPSPGSKLLIGLCEQNEVYKLSSEGKTHGSLIILAIDEGCVGR
FOODPFIITAGGAENELPLIISTQOANADLISITIDYAVSKDPHYGVYAP
KDAIISREEMLANALUTTEFSEQDMARQAEKGRMPSEPNFRNININQRYTSP
FISNSVETLCEMPTIPKRWYAGDLSARDIDALVIACEADGQWVDYFFFWTPQK

```

TLEERTKTDAPYVWVWREGILRTTPASVYAFVVAIAELTGFDLTLSNAFDWR
 DOFRKDALATLSLPEVPGOGRDNGPAVDTEFSLALNGRRHRHMHVLTLMCAVANA
 VKKAAGNRKDKSKATGRIDGVAMTMSVGAANCEVTEOGGGDFDFTFRPLSM"
 2056..3330
 /gene="3"
 2056..3330
 /gene="3"
 /note="Gp3"

```

/note="putative"
3348, .4025
/gene="4"
3348, .4025

```

```
028. .5185
gene="5"
028. .5185
```

```

ote="putative"
19.5545
ane="6"
19.5545

```

KVOBELDITAGSTRFLDLEOKLASGSEKSEVLA
 KATLEBDSQAMVULFPAOKALETSTGYKVSQLOJSDL
 KISGASNDAGSGSLIQPMQPEKSESERAAEELIKSMGKCGTGAFTPK
 DVAAERALKPESDITFPMQIINVTILAHWVQASROWADAPIMQSYREVEYENNN
 EBQOLINGGTGSDNIEGKATVATVDSLNATGPTADILAHAYOYTESBSASGVL
 LNRDWHNLALTKDNEGRYIEGSGQATSNMIMGGLVPTPKQAAGTTFVGSQPMASQ
 VMDRATVYESREDRDNFKMNLIIICERLALAHYRPAITLIKGTFSGS
 5192..5213
 /note="putative"
 5219..5545
 /gene="6"
 5219..5545
 /gene="6"
 /codon_start=1
 /transl_table=1
 /evidence=not_experimental
 /product="gp6"
 /protein_id="AA39109.1"
 /db_xref="GI:6901356"
 /translation="MADLVDTISLKFQQLIEEDRDLEITLYAQAAPGTCYRMC
 DEPMKVAADIIPAAYGAVILWFADMFHRTAQSQVLYENAAAEPMKTLHNRWGKA


```

gene      ESEBS"
          5545..5883
          /gene="7"
          5545..5883
          /gene="7"
          /note="Gp7"
          /codon_start=1
          /transl_table=1
          /evidence=not_experimental
          /product="putative head-tail adaptor"
          /protein_id="AAE31100.1"
          /db_xref="GI:6901597"
          /translation="MEPGRFRNRVKILFTTSDPSGQPVESWTGKNPVAEYKIGIS
          REOLSGAETAAQATIRVWMAFRSELNASSRLLEVLSGPYKGOVNIITGPVANNATGTRL
          ELICKGAERK"
          5880..6329
          /gene="10"
          5880..6329
          /gene="10"
          /codon_start=1
          /transl_table=1
          /evidence=not_experimental
          /product="Gp10"
          /protein_id="AAE31101.1"
          /db_xref="GI:6901598"
          /translation="MIETSLDFGSLNDIAKLEALSRANNKYLKDATRAGAVLKEE
          VIDRAPVPTGKLKKNVYVTKSRERGEISSGVHIRGVNPRGNSDNTKANNPRAAF
          YWREVELGTANMPAHFVRPAVDTRREEASVALAHNQAIDVLEK"
          6326..6673
          /gene="11"
          6326..6673
          /gene="11"
          /codon_start=1
          /transl_table=1
          /evidence=not_experimental
          /product="Gp11"
          /protein_id="AAE31102.1"
          /db_xref="GI:6901599"
          /translation="MNEDNTYALLSPPLAEGRVYPVAPLGSDDKPSVSPMIFITVD
          DVSADVLGGOAESRVSQVDVYSTIAESRLDLVLASLEPLTTEYVKIIPYEPDY
          RLYRATLDKRVTP"
          6672..6719
          /note="putative"
          /evidence=not_experimental
          6733..7437
          /gene="12"
          6733..7437
          /gene="12"
          /note="Gp12"
          /codon_start=1
          /transl_table=1
          /evidence=experimental
          /product="major tail subunit"
          /protein_id="AAE31091.1"
          /db_xref="GI:6901588"
          /translation="MSALYEKSQLTKILISLSPSTKEMDSATFLDLSCTIKETIOPTG
          GOKODIDVTTLCTEOENINGLPSSEISGNFYNNPADALDAYDNDTMYGQII
          FPGNGDFKFLAEVROHWSGNGVVAATFSLRKGPYPIDSVLKITDLDSSLSVA
          VGAISMAYVAAGKRPYATYTKKAGSTVSGTSDTFNKATAVSGDAGDITCVYDSS
          SPKVTYSACITLTIS"
          7443..7462
          /note="putative"
          7472..7858
          /gene="13"
          7472..8145
          /gene="14"
          7472..7858
          /gene="14"

terminator
gene
CDS

Query Match      45.7%: Score 426.8; DB 7: Length 39732;
Best Local Similarity 95.1%: Pred. No 1,2e-124;
Matches 451: Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 456 aacgagcgaatlaaccatcgltga-gtcaaatlaccatlttaccatgaatcaat 514

```

```

Db 35920 AACGAGTCATFTTATCCATGCTGATGTCAAAATTTCACCAACTTATTCAAAAAGCAAT 35861
QY 515 atcagccgttaataatgltgcacatcgltgcacatcatgctgtaacgltgaccgcatcc 574
Db 35860 ATCATGCGCGTTAAATATGTTGCCATCCGTGGCAATCATAGTGTGAAGGTGACCGGCTTC 35801
QY 575 aaatgtctctcgatgtgactctctctgttgacatltgcacacagagcgatcatcagc 634
Db 35800 AAAATGTTGTGTGAATGTGACTCTCTCTGTGTGCAATTGACACACAGAGGTATACAGC 35741
QY 635 ggcctaacatgctgtaacacagtggtggtglaagglttgagatlgacatcgtaacagc 694
Db 35740 GCGTTAACAATGCGGTAGCACAGGTGGGTGGGTGAGGTTTGGATTTAGCATCGTTACAGCG 35681
QY 695 cgatagctgagctgtgcatccttgaaatagcgaagcctltgacatctccgacact 754
Db 35680 CGATATGCGGCGGCTGTGCAATTTGAATAGCCGACACACTTTCATCTCCGCAATCT 35621
QY 755 ttctcgacaactctcccccaagctctgttttgcaataltcaacgcagcgactgacga 814
Db 35620 TTCTCAACAACCTCTCCCCCAGACTCTGTGTTTGGCAATATCAACCCAGGCTGTACCG 35561
QY 815 tggcaatctctgcatcttgcccccggtcgcgcgcaactagcgcaataatccgcatagcg 874
Db 35560 TGGCAATCTCTGCATCTTGGCGCCGCGGTGCGGCACTAGCGCAATAATCCGATTAAGCG 35501
QY 875 aaggttgcgagcaactgacgtacacttgcttagtatttccctcaagctgcgcc 928
Db 35500 AATGTTGCGAGCACTTGCACTTGGCTTGTAGTATTTCCTTCGAGCTTGGCC 35447

```

Search completed: January 17, 2002, 16:25:58
Job time: 13879 sec

3140 T; 0 other;

		mismatches	0;	Indels	0;	Gaps	0;
1	aaacgcgagaggttgcttaqccgcacactcctcaaac						

b
11000 aaacgcagcagcttcattgacccmactat
60

61 aaagcgccgacgcytaactattactaatgatcagaacagacaglygtcaagcttcagt 1105
|||||
1106 aaagcccccacgaatcccccttccgcggctttcacgaaggtcatgtgta 1105

121 tcaattatataatcttgcacgagcagcagcagtggtctaacgctcaatt 111

11120 tgggtcgtcgtctgctctggcgcgatgaacgcctgtacgacattgtgtgatccggctctgc 180

181 ttccggtatcgttaattcagcacacaacgaaagaggaactctc-
||||| -gaaacggcgcacccggtctctgc 111

11180 ttccggtattcgcttaattcagcacacgaagaagcactgctcaaccagctcgcgcac 240
241 tcttcacgattatfcacatcaaatgataaaagtacggcttgcacacacacacacac

11240 tcttcacgaattatgaatgagctggtgtgcagatatataaaatccgaac 300

301 cgtatgcagcgtctactatttaactgcgaactggttcggatgtcatttttcagacgc 360
|||||
11300 cgtatgcagcgtctactatttaactgcgaactggttcggatgtcatttttcagacgc 360
|||||

21500 cytlatgcagcgtctacatactacacctgcgaactgtttcggatgcattttcaaacaccc 580

361 tctgctgcatggttgagttccagacatacgtcgaagtacacactagcgaatcg 420
|||||
11360 tctgctgcatggttgagttccagacatacgtcgaagtacacactagcgaatcg 420

421 gtagtaagcgccgcctcttcttcatctaatat
11419

481 gtcgaatttaccatcctccacgagcgaaatcaaccatcgttga 11479
11420 gttagtaagcgcgctctttcatctcacccaagaagcgatcaaccatcgttga 480

11480tatttcaataagtcataatcatgcggtcaatatgttgcaccc 510
ataa.....

541 gtggcaatcatgctgctaagctgtgaacgcgcatlGaaatatttat
gccaadattaccaaatltaattcaataagtcataatcatgcccgttaataatgttcgcatcc 11539

11540 gtggcaatcatgtcgtctaacgtgttgacgcacatfcaactc 600

601 ttgtgtgcattgcaccaccagagcgtcatatacagcgtctaacaagtgcgtaccacagttgg 660

XX DNA; 11933 BP.

04-SEP-2001 (first entry)

Human factor VIII expressing rAAV vector pAAV-F8-1 partial sequence

PAAV-F8-1; ds. X

Chimeric - Adeno associated virus.
Chimeric - Mus sp.
Chimeric - Homo sapiens

Chimeric - *Oryctolagus cuniculus*.

MO200145510-A1.
28-JUN-2001.

21-DEC-2000:

22-DEC-1999; 99US-0470618.

(AVIG-) AVIGEN INC.

Couto LB, Colosi PC, Qian X;
WPI; 2001-417955/44.

Treating blood clots

claim 18. For example, administering recombinant adeno-associated vectors which express blood coagulation factor VIII -

supp; English.

The present invention relates to a method for treating a subject suffering from a blood clotting disorder. The method comprises administering a recombinant adeno-associated virus encoding a nucleotide sequence encoding a factor VIII protein.

second recombinant adeno-associated virus (rAAV) comprising the sequence encoding the light chain of factor VIII and a sequence encoding the heavy chain of factor VIII.

expression of biological functions. The rAAV vector provides high level

CC backbone. It comprises HNF-3, mouse albumin promoter, a synthetic intron
CC elongation factor 1alpha (EF1alpha) and immunoglobulin G
CC based on human elongation factor 1alpha (EF1alpha) and IAV coding
CC (1993) intron sequences. B-domain deleted beta-globin sequence and the
CC sequence, poly A signal based on rabbit beta-globin sequence and both
CC inverted terminal repeats (ITRs) at the ends. The vector encodes both
CC inverted terminal halves of human factor VIII.

Sequence	11933 bp;	3258 A;	2818 C;	2717 G;	3140 T;	0 other;
CC						
XX						
SO						

	Score	934; Du	Indels	Gaps	
Query Match	100.0%	Pred. No. 4.9e-312;			
Best Local Similarity	100.0%;	Mismatches	0;		
Matches	934;	Conservative			

[illegible]

	901	tgcctagatattccttcaagctgccccttcagg	934
QY			
	11900	tgccttagatattccttcaagctgccccttcagg	11933
Db			

RESULT 3
AAAI1992
11003 standard; DNA; 37856 BP

ID	AAAl1992	07-AUG-2000	(first entry)
XX			
AC	AAAl1992		
XX			

XX S. celluloseum DNA encoding polyketide and heteropolysaccharide biosynthetic genes
DE
XX Polyketide biosynthesis; heteropolysaccharide biosynthesis; mutasynthesis
XX epothilone; cytotoxic; immunosuppressant; antibiotic; antifungal
KM
KM plant-protection; ds.
XX
XX *Saccharatum celluloseum*.

OS	Location/Qualifiers
XX	complement (3398..6100
FH	
Key	
305	

```

FT      CDS      /*tag= a
FT      /product= "ORF1-trNA synthetase
FT      /note= "gtg start codon"
FT      complement (6374..7111)

```

FT	CDS	/*tag=	b
FT		/product=	"ORF2-monoxygenase
FT		complement	(8433..9550)
FT		/*tag=	c
CDS			transfer

FT	/product= "ORF3-aminocyclase"
FT	/note= "AGT start codon given in the specification"
FT	9855..11393
CDS	join(ORF3-decarboxylase"

```

FT      /tag      "ORF4-tyrosine/pos=
FT      /product= "GTG start codon"
FT      /note= "13658
FT      12212..13658

```

CDC	/*tag=	C	ORF5-3-oxoacyl-ACF	15374..19984
FT			/product=	"ACC start codon"
FT			/note=	"ACC start codon"
FT				15374..19984

```

CDS      /*tag= 1 "ORF6-polyketide synthase"
FT
FT      /product= "ORF6-polyketide synthase"
FT      20003..27889
FT      /*tag= 9 "nort-nentide synthetase"
FT

```

FT	/product= "Orf1 / Pef-
FT	28251..29400
FT	CD5
FT	/*tag= h
FT	/product= "Orf8-transpeptidase"
	21720

FT	CDS	complement (30040...32120)
FT		/tag= 1
FT		/product= "ORF9-regulation element"
FT		/note= "CGC stop codon"

Accession	Gene	Product	Length (aa)	Mass (kDa)	pI	Ref
U00096	oriC	oriC	245	26.5	4.5	[1]
U00096	oriT	oriT	245	26.5	4.5	[1]
U00096	oriV	oriV	245	26.5	4.5	[1]
U00096	oriW	oriW	245	26.5	4.5	[1]
U00096	oriX	oriX	245	26.5	4.5	[1]
U00096	oriY	oriY	245	26.5	4.5	[1]
U00096	oriZ	oriZ	245	26.5	4.5	[1]
U00096	oriAA	oriAA	245	26.5	4.5	[1]
U00096	oriAB	oriAB	245	26.5	4.5	[1]
U00096	oriAC	oriAC	245	26.5	4.5	[1]
U00096	oriAD	oriAD	245	26.5	4.5	[1]
U00096	oriAE	oriAE	245	26.5	4.5	[1]
U00096	oriAF	oriAF	245	26.5	4.5	[1]
U00096	oriAG	oriAG	245	26.5	4.5	[1]
U00096	oriAH	oriAH	245	26.5	4.5	[1]
U00096	oriAI	oriAI	245	26.5	4.5	[1]
U00096	oriAJ	oriAJ	245	26.5	4.5	[1]
U00096	oriAK	oriAK	245	26.5	4.5	[1]
U00096	oriAL	oriAL	245	26.5	4.5	[1]
U00096	oriAM	oriAM	245	26.5	4.5	[1]
U00096	oriAN	oriAN	245	26.5	4.5	[1]
U00096	oriAO	oriAO	245	26.5	4.5	[1]
U00096	oriAP	oriAP	245	26.5	4.5	[1]
U00096	oriAQ	oriAQ	245	26.5	4.5	[1]
U00096	oriAR	oriAR	245	26.5	4.5	[1]
U00096	oriAS	oriAS	245	26.5	4.5	[1]
U00096	oriAT	oriAT	245	26.5	4.5	[1]
U00096	oriAU	oriAU	245	26.5	4.5	[1]
U00096	oriAV	oriAV	245	26.5	4.5	[1]
U00096	oriAW	oriAW	245	26.5	4.5	[1]
U00096	oriAX	oriAX	245	26.5	4.5	[1]
U00096	oriAY	oriAY	245	26.5	4.5	[1]
U00096	oriAZ	oriAZ	245	26.5	4.5	[1]
U00096	oriBA	oriBA	245	26.5	4.5	[1]
U00096	oriBB	oriBB	245	26.5	4.5	[1]
U00096	oriBC	oriBC	245	26.5	4.5	[1]
U00096	oriBD	oriBD	245	26.5	4.5	[1]
U00096	oriBE	oriBE	245	26.5	4.5	[1]
U00096	oriBF	oriBF	245	26.5	4.5	[1]
U00096	oriBG	oriBG	245	26.5	4.5	[1]
U00096	oriBH	oriBH	245	26.5	4.5	[1]
U00096	oriBI	oriBI	245	26.5	4.5	[1]
U00096	oriBJ	oriBJ	245	26.5	4.5	[1]
U00096	oriBK	oriBK	245	26.5	4.5	[1]
U00096	oriBL	oriBL	245	26.5	4.5	[1]
U00096	oriBM	oriBM	245	26.5	4.5	[1]
U00096	oriBN	oriBN	245	26.5	4.5	[1]
U00096	oriBO	oriBO	245	26.5	4.5	[1]
U00096	oriBP	oriBP	245	26.5	4.5	[1]
U00096	oriBQ	oriBQ	245	26.5	4.5	[1]
U00096	oriBR	oriBR	245	26.5	4.5	[1]
U00096	oriBS	oriBS	245	26.5	4.5	[1]
U00096	oriBT	oriBT	245	26.5	4.5	[1]
U00096	oriBU	oriBU	245	26.5	4.5	[1]
U00096	oriBV	oriBV	245	26.5	4.5	[1]
U00096	oriBW	oriBW	245	26.5	4.5	[1]
U00096	oriBX	oriBX	245	26.5	4.5	[1]
U00096	oriBY	ori				

E1	33126..
CDS	/start= k
FT	/product= "ORF1-regulation element"
FT	/note= "GTG start codon"
FT	/note= 34077

FT	33661..3407,
FT	/ *tag= 1
FT	/product= "ORF12-regulation element"
FT	complement (35255..35616)
FT	

FT	CDS
FT	/*tag= m
FT	/product= "ORF13-transcription regulator
FT	complement (35730..36242)
FT	/*tag= n
CDS	transcription regulator

PT /product="OKR14-111111"
PT /note="GIG start codon"
PT
XX

PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0126785.
PR	06-APR-1999;	99US-0127462.
PR	08-APR-1999;	99US-0128234.
PR	15-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129645.
PR	21-APR-1999;	99US-0130077.
PR	23-APR-1999;	99US-0130449.
PR	28-APR-1999;	99US-0130510.
PR	28-APR-1999;	99US-0132468.
PR	30-APR-1999;	99US-0132468.
PR	04-MAY-1999;	99US-0133407.
PR	05-MAY-1999;	99US-0133468.
PR	06-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132485.
PR	07-MAY-1999;	99US-0132487.
PR	11-MAY-1999;	99US-0132463.
PR	14-MAY-1999;	99US-0134356.
PR	14-MAY-1999;	99US-0134318.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134964.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135352.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136031.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137322.
PR	03-JUN-1999;	99US-0137322.
PR	04-JUN-1999;	99US-0137328.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138540.
PR	14-JUN-1999;	99US-0139884.
PR	16-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139432.
PR	17-JUN-1999;	99US-0139432.
PR	18-JUN-1999;	99US-0139443.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139465.
PR	18-JUN-1999;	99US-0139750.
PR	21-JUN-1999;	99US-0139753.
PR	22-JUN-1999;	99US-0139817.
PR	23-JUN-1999;	99US-0139867.
PR	23-JUN-1999;	99US-0140353.
PR	24-JUN-1999;	99US-0140353.
PR	24-JUN-1999;	99US-0140355.
PR	29-JUN-1999;	99US-0140823.
PR	30-JUN-1999;	99US-0140991.
PR	01-JUL-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141847.
PR	02-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	08-JUL-1999;	99US-0142303.
PR	09-JUL-1999;	99US-0142803.
PR	12-JUL-1999;	99US-0142950.
PR	13-JUL-1999;	99US-0142977.
PR	15-JUL-1999;	99US-0143542.
PR	15-JUL-1999;	99US-0143624.
PR	16-JUL-1999;	99US-0144008.

us-09-740-211-13_copy-11000-11933.rng

```

08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.5%; Score 33; DB 21; Length 1063;
Best Local Similarity 52.6%; Pred. No. 0.88; Gaps 0
Matches 72; Conservative 0; Mismatches 65; Indels 0

OY 466 ttaaccacgtgtgacaaatttaacccaatttttaataaagtcataatcagtcgt 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 ttccaacatcattgatatgatgtgaactatattgaacaacacagcatcctctct 925
OY 526 aatagtttccatccgtgagcatcatctgtctaagctgtgacggatccaatgtgtc 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 tgaagatgtcatcgcgaataagaacagtgcaacatttgatctctctttgtttg 985
OY 586 tgcagatgactctctct 602
    ||| ||| ||| ||| |||
Db 986 tgccttatgtccttcat 1002
    ||| ||| ||| ||| |||

RESULT 5
AAF28541 standard; DNA; 49617 BP.
ID AAF28541
XX AAF28541;
AC AAF28541;
XX 04-APR-2001 (first entry)
DT
XX Genomic fragment #28.
DE
XX Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KM bronchopulmonary; endocarditis; meningitis; ss.
XX Moraxella catarrhalis.
OS
XX WO200078968-A2.
XX
XX 28-DEC-2000.
PD
XX 16-JUN-2000; 2000WO-US16649.
PF
XX 18-JUN-1999; 99US-0140121.
PR
XX (INCY-) INCYTE GENOMICS INC.
PA
XX

```



```

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25: SEQ ID No 2751: 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SNP). The present sequence is one such probe. The SNPs are derived
CC from human Hela cells. The SNPs can be used to produce a single exon
CC microarray which can be used for measuring human gene expression
CC sample derived from human cervical epithelial cells. By measuring in a
CC of diseases of the cervix, the probes are therefore useful in grading and/or staging
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp://pub.int/pub/published_pcr_sequences.
XX
SQ Sequence 469 BP: 125 A; 159 C; 88 G; 97 T; 0 other.

Query Match
Best Local Similarity 3.4%; Score 31.8; DB 22; Length 469;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 98 cagacagatggcagcgctcagcttgatggatgctgctgagcgagcagcctgta 157
Db ||||||| || || || || || || || || || || || || || || || ||
OY 115 CAGACGAGAGCAGAGATTGAGTGGGGTCTGCTGCTTGGAGGACAGTGGCAGATC 157
Db ||||||| || || || || || || || || || || || || || || || ||
OY 158 ccatttgatccggctgctgcttcgcgta 188
Db ||||||| || || || || || || || || || || || || || || || ||
55 AAACCTTGGCTTCTGCGAGAGCCTTCTCGCA 25

RESULT 7
AA134173.C
ID AA134173 standard; DNA; 469 BP.
XX
AC AA134173;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #2859 used to measure gene expression in human placenta sample.
XX
KW Probe: microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001, 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
(MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX

```


[illegible][illegible]

[illegible]

Search completed: January 17, 2002, 16:33:41
Job time: 12526 sec

Run on: January 17, 2002, 11:00 AM
(without alignments)
1789.295 Million cell updates/sec

NS-09-740-211-13_COPY_11000_11933

```

Title: 02
perfect score: 934
1 aaacccagaagatttttag.....cttcaagctgcacctgcagg 934

```

Scoring table:

	IDENTITY_NUC
Gapop 10.0 ,	Gapext 1.0

Searched: 351203 seqs, 115250000

Total number of hits satisfying chosen parameter.....

Minimum	DB seq	length:	2000000000
Maximum	DB seq		

```

post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```
Database : Issued_Patents.mdb
1: /cgn2-6/plodata/2/lna/5A-COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	934	100.0	11933	4	US-09-470-618-13	Sequence 13, Appl
2	934	100.0	11933	4	US-09-364-862-13	Sequence 13, Appl
3	37.6	4.0	7218	4	US-08-232-463-14	Sequence 14, Appl
4	31.6	3.4	4668	4	US-09-045-301-1	Sequence 1, Appl
5	31.2	3.3	2940	4	US-09-195-868-11	Sequence 11, Appl
6	31.2	3.3	3691	3	US-09-195-868-12	Sequence 12, Appl
7	31.2	3.3	4147	3	US-08-360-005-1	Sequence 1, Appl
8	31.2	3.3	4147	3	US-09-418-540-1	Sequence 7, Appl
9	31.2	3.3	4870	4	US-08-664-962B-7	Sequence 7, Appl
10	31.2	3.3	4870	4	US-09-311-743-7	Sequence 17, Appl
11	31	3.3	289	4	US-09-007-005-17	Sequence 17, Appl
12	31	3.3	289	4	US-09-244-796-17	Sequence 17, Appl
13	30.8	3.3	1951	1	US-08-487-890A-112	Sequence 112, App
14	30.8	3.3	1951	1	US-08-487-835-112	Sequence 112, App
15	30.8	3.3	1951	2	US-08-337-483-112	Sequence 112, App
16	30.8	3.3	1951	3	US-08-478-473-112	Sequence 112, App
17	30.8	3.3	1951	3	US-08-474-671-112	Sequence 112, App
18	30.8	3.3	1951	3	US-08-483-577A-112	Sequence 112, App
19	30.8	3.3	1951	4	US-08-998-416-514	Sequence 514, App
20	29.6	3.2	4253	4	US-08-577-483-7	Sequence 7, Appl
21	29.6	3.2	4253	4	US-08-443-639-7	Sequence 7, Appl
22	29.6	3.2	4253	4	US-08-443-994-1	Sequence 4, Appl
23	28.8	3.1	1566	3	US-08-945-994-4	Sequence 114, App
24	28.6	3.1	1566	3	US-08-487-890A-114	Sequence 114, App
25	28.6	3.1	1955	1	US-08-476-435-114	Sequence 114, App
26	28.6	3.1	1955	2	US-08-337-483-114	Sequence 114, App
27	28.6	3.1	1955	2		

C 28	28.6	3.1	1955	2	US-08-478-373-114	Sequence 114, App
C 29	28.6	3.1	1955	3	US-08-474-671-114	Sequence 114, App
C 30	28.6	3.1	1955	4	US-08-483-577A-114	Sequence 114, App
C 31	28.6	3.1	1955	4	US-08-897-438-114	Sequence 114, App
C 32	28.4	3.0	585	2	US-08-882-704A-1	Sequence 4, Appl
C 33	28.4	3.0	7742	2	US-08-882-704A-4	Sequence 2, Appl
C 34	28.4	3.0	4403765	4	US-09-103-840A-2	Sequence 1, Appl
C 35	28.4	3.0	4411529	4	US-09-103-840A-1	Sequence 87, Appl
C 36	28	3.0	1142	1	US-08-105-483-87	Sequence 87, Appl
C 37	28	3.0	1142	1	US-08-709-209-87	Sequence 87, Appl
C 38	28	3.0	1142	1	US-08-458-101-87	Sequence 1, Appl
C 39	28	3.0	2656	1	US-07-961-522-1	Sequence 1, Appl
C 40	28	3.0	2656	1	US-08-217-438-1	Sequence 1, Appl
C 41	28	3.0	2656	1	US-08-321-978-1	Sequence 1, Appl
C 42	28	3.0	2656	2	US-08-710-584-1	Sequence 1, Appl
C 43	28	3.0	2656	1	US-07-984-044A-3	Sequence 3, Appl
C 44	28	3.0	3058	1	US-08-458-393-3	Sequence 7, Appl
C 45	27.8	3.0	405	3	US-08-950-720A-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-470-618-13
US/09470618

Sequence 13, Appl. No. 6200560

GENERAL INFORMATION:
APPLICANT: Couto, Linda B.

APPLICANT: Colosi, Peter
TITLE OF INVENTION: Adeno-

TITLE OF INVENTION: by Target Cell
; TITLE OF INVENTION: Avigen-04082
; REFERENCE: Avigen-04082

FILE REFERENCE NUMBER: US/09/4/U, b1
CURRENT APPLICATION NUMBER:
ISSUING DATE: 1999-12-22

CURRENT FILING DATE: 09/364,862
EARLIER APPLICATION NUMBER: 1999-07-30

EARLIER FILING DATE: 1999-03-24
EARLIER APPLICATION NUMBER: 60/125,974

EARLIER FILING DATE: 1999 05 24
; EARLIER APPLICATION NUMBER: 60/1
; EARLIER APPLICATION NUMBER: 60/10-30

EARLIER FILING DATE: 1998-10-15
NUMBER OF SEO ID NOS: 15

NUMBER OF: PatentIn Ver. 2.0
SOFTWARE: PatentIn Ver. 2.0
FILE NO: 13

; SEQ ID NO 1
; LENGTH: 11933
; DRY

```

; TYPE: DNA
; ORGANISM: Artificial Sequence

```

FEATURE:	Description of Artificial Sequence:
OTHER INFORMATION:	

US-09-470-618-13

Query Match	Score	DB	Length	Cons
100.0%	934	4	1133	
100.0%	pred	NO	0	

Best Local Similarity	100.0%	0;
Best Local Similarity	0;	Inders
Matches 934; Conservative	0;	Mismatches

1 aaacacgaagatgtttagcgcgaacctctgccaaccgcttcaacgaagtcattgtga 60

[illegible]

Dd 11000 aaacgycagcaggatctgaacttctaataattactaatgcagacagacagltgctacgcgtcagt 1

61 aaagggccgcacgcaaa
|||
qy |||attactaatgaatcagacagacagtgtctacgcttcagtt

Db 11060 aaagccgcgacgclaacatcaccaaa;

QY 121 tgggtctgctgctgcacggcagcgacga-
|||||
|||atagcccttacqcatlctgtgatccggttcgc 1

[illegible]

181 ttccggtatcgccttaattcagcacacagcyaaagagcacaacg3
QY

Db 11180 ttccggtattcgttaattcaccaacgcgaagaagatcgccaaa--

OY 241 tcttcacgattatcgcactcaatgtctcttaccctgtttgcagaca

US-09-364-862-13

Query Match 100.0%; Score 934; DB 4; Length 11933;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 934; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 11240 tctcaagatlatcgactcaatgctctaccgtctgctgagatataaaatccgaac 11299
QY 301 cgtatgacgagcttaactactactgcaactgttccgagattgcaactc 360
Db 11300 cgtatgacgagcttaactactactgcaactgttccgagattgcaactc 11359
QY 361 tctgctgagatggttggagttccagagatcgctgcaagtgaacactagcggaacg 420
Db 11360 tctgctgagatggttggagttccagagatcgctgcaagtgaacactagcggaacg 420
QY 421 gtagtaagcgccgctcttcttctcactaccacacagagaaattacccatctga 480
Db 11420 gtagtaagcgccgctcttcttctcactaccacacagagaaattacccatctga 480
QY 481 gtaaatattacccaatlttataatgaatgaatcagcgttaatatgttgcacac 540
Db 11480 gtaaatattacccaatlttataatgaatgaatcagcgttaatatgttgcacac 540
QY 541 gtagcaatcatgctgttaacgctgtaacgcaatcaaaatgttctgctgcaattc 600
Db 11540 gtagcaatcatgctgttaacgctgtaacgcaatcaaaatgttctgctgcaattc 600
QY 601 ttgttgacatgacccacacagagcgtaacagcgcttaacagtgctgacacagtg 660
Db 11600 ttgttgacatgacccacacagagcgtaacagcgcttaacagtgctgacacagtg 660
QY 661 ttgggttaaggttggagttgagatgacatcgtaacagcgatagctgctgcaactc 720
Db 11660 ttgggttaaggttggagttgagatgacatcgtaacagcgatagctgctgcaactc 720
QY 721 tgaatagcgagcgcttctgcatcttcgcacatcttctcgaacatcccccaagctc 780
Db 11720 tgaatagcgagcgcttctgcatcttcgcacatcttctcgaacatcccccaagctc 780
QY 781 tgtttggcaatatcaaccgacgagcgttaccatgagaaatctctgacatctgccc 840
Db 11780 tgtttggcaatatcaaccgacgagcgttaccatgagaaatctctgacatctgccc 840
QY 841 cgtcgagcaatacggcaataatcgcatgaaggaatgttggagacactgagatcct 900
Db 11840 cgtcgagcaatacggcaataatcgcatgaaggaatgttggagacactgagatcct 900
QY 901 tgccttagatttctcctcaagctgcccctgcaag 934
Db 11900 tgccttagatttctcctcaagctgcccctgcaag 11933

```

RESULT 2
 US-09-364-862-13

Sequence 13, Application US/09364862
 Patent No. 6221349

GENERAL INFORMATION:

APPLICANT: Coultio, Linda B.

APPLICANT: Coultio, Peter C.

TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII

TITLE OF INVENTION: BY TARGET

FILE REFERENCE: AVIGEN-03743

CURRENT APPLICATION NUMBER: US/09/364,862

EARLIER APPLICATION NUMBER: 60/125,974

EARLIER FILING DATE: 1999-03-24

EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 13

LENGTH: 11933

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Description of Artificial Sequence: synthetic

RESULT 3

```

QY 1 aacagcagagaggttggtagcgacactctcgcacccgcttaccagagtgatgta 60
Db 11000 aacagcagagaggttggtagcgacactctcgcacccgcttaccagagtgatgta 60
QY 61 aaagcgcagcgcttaactactactaataatcaagacagagtggtcagcgtaagt 120
Db 11060 aaagcgcagcgcttaactactactaataatcaagacagagtggtcagcgtaagt 120
QY 121 tgggtgtcgtcttctcgtggcgagatgacgctgtacgatttggtccggtctgc 180
Db 11120 tgggtgtcgtcttctcgtggcgagatgacgctgtacgatttggtccggtctgc 180
QY 181 ttcggatctcgttaattcagcacaacaggaagagcacttggtacacagcgctgc 240
Db 11180 ttcggatctcgttaattcagcacaacaggaagagcacttggtacacagcgctgc 240
QY 241 tcttacgattatgactcaatgctcttactcgtgtgtgagatataaaatccgaac 300
Db 11240 tcttacgattatgactcaatgctcttactcgtgtgtgagatataaaatccgaac 300
QY 301 cgtatgacgagcttaactactactcctcggaactgttggagattgcaatttgcagac 360
Db 11300 cgtatgacgagcttaactactactcctcggaactgttggagattgcaatttgcagac 360
QY 361 tctgctgagatggttggagttcgcagacgagatgcaagtgacacacagagagatg 420
Db 11360 tctgctgagatggttggagttcgcagacgagatgcaagtgacacacagagagatg 420
QY 421 gtagtaagcgccgctcttctcactacacacagcggaatlttaacacatctgta 480
Db 11420 gtagtaagcgccgctcttctcactacacacagcggaatlttaacacatctgta 480
QY 481 gtaaatattacccaatlttataatgaatgaatcagcgttaatatgttgcacac 540
Db 11480 gtaaatattacccaatlttataatgaatgaatcagcgttaatatgttgcacac 540
QY 541 gtagcaatcatgctgttaacgctgtaacgcaatcaaaatgttctgctgcaattc 600
Db 11540 gtagcaatcatgctgttaacgctgtaacgcaatcaaaatgttctgctgcaattc 600
QY 601 ttgttgacatgacccacacagagcgtaacagcgcttaacagtgctgacacagtg 660
Db 11600 ttgttgacatgacccacacagagcgtaacagcgcttaacagtgctgacacagtg 660
QY 661 ttgggttaaggttggagttgagatgacatcgtaacagcgatagctgctgcaactc 720
Db 11660 ttgggttaaggttggagttgagatgacatcgtaacagcgatagctgctgcaactc 720
QY 721 tgaatagcgagcgcttctgcatcttcgcacatcttctcgaacatcccccaagctc 780
Db 11720 tgaatagcgagcgcttctgcatcttcgcacatcttctcgaacatcccccaagctc 780
QY 781 tgtttggcaatatcaaccgacgagcgttaccatgagaaatctctgacatctgccc 840
Db 11780 tgtttggcaatatcaaccgacgagcgttaccatgagaaatctctgacatctgccc 840
QY 841 cgtcgagcaatacggcaataatcgcatgaaggaatgttggagacactgagatcct 900
Db 11840 cgtcgagcaatacggcaataatcgcatgaaggaatgttggagacactgagatcct 900
QY 901 tgccttagatttctcctcaagctgcccctgcaag 934
Db 11900 tgccttagatttctcctcaagctgcccctgcaag 11933

```

```
US-08-232-463-14
? US-08-232-463-14 Application US/08232463
? Patent No. 5670367
? GENERAL INFORMATION:
? APPLICANT: DORNER, F.
? APPLICANT: SCHEFLINGER, F.
? APPLICANT: FALKNER, F. G.
? TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
? NUMBER OF SEQUENCES: 52
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Foley & Lardner
? STREET: 1800 Diagonal Road, Suite 500
? CITY: Alexandria
? STATE: VA
? COUNTRY: USA
? ZIP: 22313-0299
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/232,463
? FILING DATE:
? CLASSIFICATION: A35
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/07/935,313
? FILING DATE: EP 91 114 300.6
? APPLICATION NUMBER: EP 91 114 300.6
? FILING DATE: 26-AUG-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: BENT, Stephen A.
? REGISTRATION NUMBER: 29,768
? REFERENCE/DOCKET NUMBER: 30472/114 IMMU
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703)836-9300
? TELEFAX: (703)683-4109
? TELEX: 899149
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 7218 base pairs
? type: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? IMMEDIATE SOURCE:
? CLONE: pTZ9pt-P1S
? US-08-232-463-14
Query Match 4.0%; Score 37.6; DB 1; Length 7218;
Best Local Similarity 2.7%; Pred. NO. 0.015;
Matches 7; Conservative 152; Mismatches 101; Indels 0; Gaps 0.
```

```

RESULT 4
US-09-045-301-1
Sequence 1, Application US/09045301A
Patent No. 6265388
GENERAL INFORMATION:
APPLICANT: Petri, James W.
APPLICANT: Olson, Karen A.
TITLE OF INVENTION: Antisense Inhibition of Angiogenin Expression
FILE REFERENCE: 10498/05286
CURRENT APPLICATION NUMBER: US/09/045.301A
CURRENT FILING DATE: 1998-03-20
EARLIER APPLICATION NUMBER: 60/041182
EARLIER FILING DATE: 1997-03-21
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 4668
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1809)..(2252)
US-09-045-301-1

```

```

Query Match          3.4%; Score 31.6; DB 4; Length 4668;
Best Local Similarity 53.2%; Pred. No. 1.4; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 59;

QY   378 gatttcacagacatcagctcgaaagtaccgaactaggcggaatcgtagtaagaagccgcctc 437
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   4525 gaocgccgcatgcatcagcactgcagcgtggaacgatcttctagtccttttglttgycctc 4584
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY   438 ttcttatcctcactcaacacacggacggacaltaacccatcgtttgagtgacaatttaccaat 497
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   4585 ttgtctcacacacccaacgaagctgtttaataataaacgctaagtcacggacggtcatt 4644
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY   498 ttatic 503
      ||| | |
Db   4645 ttatcc 4650

RESULT 5
US-09-195-868-11/c
; Sequence 11, Application US/09195868
; Patent No. 6090621
GENERAL INFORMATION:
APPLICANT: KAVANAUGH MD, MICHAEL
APPLICANT: POT PH.D., DAVID
APPLICANT: WILLIAMS MDPHD, LEWIS T.
TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
TITLE OF INVENTION: 5-PHOSPHATASES (SLPS)
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: 4560 HORTON STREET
CITY: EMERYVILLE
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/195,868
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: US/09/195,868
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
```

NAME: FIRESTONE, LEIGH H.
REGISTRATION NUMBER: 36,831
REFERENCE/DOCKET NUMBER: 1182.004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-923-2707
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2940 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-195-868-11

Query Match
Best Local Similarity 3.3%; Score 31.2; DB 3; Length 2940;
Matches 45; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 549 catgctgctaagctgtgacgcgattcaaatgtgtctgcgattgactctcttctgtgac 608
Db 1641 CAAGCAGCTCGAGTGAACCTCCAGTAGAATTGGTCTGGGACTTGTCTCAATGTGCG 1582
Oy 609 attgcacc 616
Db 1581 ATAGCACC 1574

RESULT 6

US-09-195-868-12/c
Sequence 12, Application US/09195868
Patent No. 6090621
GENERAL INFORMATION:
APPLICANT: KAVANAUGH MD, MICHAEL
APPLICANT: POT PH.D., DAVID
TITLE OF INVENTION: WILLIAMS MDPHD, LEWIS T.
TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: 4560 HORTON STREET
CITY: EMERYVILLE
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/195,868
CLASSIFICATION:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FIRESTONE, LEIGH H.
REGISTRATION NUMBER: 36,831
REFERENCE/DOCKET NUMBER: 1182.004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-923-2707
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3691 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-195-868-12

Query Match
Best Local Similarity 3.3%; Score 31.2; DB 3; Length 3691;
Matches 45; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 549 catgctgctaagctgtgacgcgattcaaatgtgtctgcgattgactctcttctgtgac 608
Db 2392 CAAGCAGCTCGAGTGAACCTCCAGTAGAATTGGTCTGGGACTTGTCTCAATGTGCG 2333
Oy 609 attgcacc 616
Db 2332 ATAGCACC 2325

RESULT 7

US-08-560-005-1/c
Sequence 1, Application US/08560005
Patent No. 6001354
GENERAL INFORMATION:
APPLICANT: Pot, David A.
APPLICANT: Williams, Lewis T.
APPLICANT: Jefferson, Anne Bennett
TITLE OF INVENTION: No. 6001354e1 Grb2 Associating Protein and Nucleic
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,005
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 2307K-0624000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4147 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 17..2944
US-08-560-005-1

Query Match
Best Local Similarity 3.3%; Score 31.2; DB 3; Length 4147;
Matches 45; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Oy 549 catgctgctaagctgtgacgcgattcaaatgtgtctgcgattgactctcttctgtgac 608
Db 1648 CAAGCAGCTCGAGTGAACCTCCAGTAGAATTGGTCTGGGACTTGTCTCAATGTGCG 1589
Oy 609 attgcacc 616
Db 11111111

1588 ATAGCACC 1581

GENERAL INFORMATION:
APPLICANT: Krystal, Gerald

COMMITTEE
MEDICAL
COMMISSION

us-09-740-211-13_copy-11000_11933.rtf

	Seqs	Mismatches	Indels	Gaps
OY	705	59	119	0
Db	247	1	0	0
OY	765	1	0	0
Db	187	1	0	0
OY	825	1	0	0
Db	127	1	0	0

[illegible]

RESULT 15

US-08-337-483-112/C
Sequence 112, Application US/08337483
Patent No. 5922562

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-ping
APPLICANT: Mordin, Andrew
APPLICANT: Klein, Michel
TITLE OF INVENTION: Transferrin Receptor Genes
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/337,483
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 1951 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
LOCATION: 1..1890
US-08-337-483-112

Query Match

Best Local Similarity 3.3%; Score 30.8; DB 2; Length 1951;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

0Y 777 gctcgttlttggaatataccgcagcgccgtgtacatggaatcctgcatctgccc 836
||| | ||||| | | ||| | | | ||| | | | ||| | | | ||| | | | |||
DB 985 GCTGTTCTTTGGCACTAAATATCCCAAAACTTTTATCGCTACTATAAAATTTCC 926
||| | ||||| | | ||| | | | ||| | | | ||| | | | ||| | | | |||
0Y 837 ccggcgctcgagcactacggaataatccgcaataagcaatgttc 882
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 925 CTAGTCTTCAGCATTAGGCCCATAAACGACCTTCTAATGTTCC 880

Search completed: January 17, 2002, 16:27:13
Job time: 12343 sec

us-09-740-211-13_copy_11000_11933.rst

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

US-09-740-211-13_COPY_11000_11933

Searched: 1135193 / segs, 33/2000000

```

post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing First 45 summaries

```

```

1:  EST.*
2:  em_estfun:*
3:  em_esthm:*
4:  em_estlin:*
5:  em_estpl:*
6:  em_esttbl:*
7:  em_esttbl:*
8:  em_esttbl:*
9:  em_hlc:*
10:  gp_est1:*
11:  gp_est2:*
12:  gp_hlc:*
13:  gp_gss:*
14:  gp_gss_fun:*
15:  em_gss_hum:*
16:  em_gss_hum:*
17:  em_gss_pln:*
18:  em_gss_pro:*
19:  em_gss_rtd:*
20:  em_gss_vrt:*
21:  em_gss_other:*

```

SUMMARIES

C	13	373	39.9	373	11	BC3124470
C	14	371.8	39.8	498	10	AA681579
C	15	360	38.5	399	10	AV749448
C	16	338.2	36.0	376	10	BE344444
C	17	327	35.2	329	10	BE334711
C	18	316.2	33.9	644	11	B1268701
C	19	305.2	32.7	408	11	B1268709
C	20	298	31.9	505	10	AW619942
C	21	294.6	31.5	558	11	B1268787
C	22	286.4	30.7	606	11	B1268691
C	23	284.2	30.3	666	11	B1268418
C	24	283.4	30.3	501	10	BE455797
C	25	260.6	27.9	291	11	N84801
C	26	246	26.3	246	11	BC455413
C	27	226.6	24.3	248	11	B1268781
C	28	22.6	23.3	425	11	BF634675
C	29	215	23.3	215	11	BF634707
C	30	184.4	19.7	211	11	BC312975
C	31	175	18.7	175	10	AV040308
C	32	154.4	16.5	242	10	AI498370
C	33	128.2	13.5	479	10	A1902016
C	34	126.4	13.5	193	11	C22175
C	35	121.8	12.5	269	11	C22255
C	36	117.2	12.5	401	11	BF657675
C	37	110.4	11.8	231	11	B1190478
C	38	94	10.1	412	11	BF529377
C	39	87	9.3	393	11	AW080651
C	40	87	9.3	121	11	BG451510
C	41	83.2	8.9	518	11	BG102530
C	42	83.2	8.9	578	11	B6035844
C	43	74	7.9	508	11	CNS01074
C	44	74	7.2	576	11	B1263666
C	45	61.2	6.6	654	11	B1265924

BG312470	MB2DAA8_F
A9A81570	MB3DAA8_F
AV799428	AV794428
BE343444	EST045522
BE353444	EST333848
BI268701	NF023B08G
AM615942	8027_MARC
BI268787	NF015C09G
BI268699	NF002403G
BI268618	NF004A30G
BE588797	EST414108
NB4801	J1047F_HOM
BC455943	M799195TM
BI268786	NF015802G
BF935475	PC4_NTL10
BF634207	NF074G04D
BG312975	WHEP412.D
AV404308	AV404308
AI1898370	EST267813
AI490216	EST218194
C22175	C22175_M19A
C22255	C22255_M19A
BF65675	0225_20_DO
BI190428	id11f5s_r
BF55927	U1-R-C2-n
AM080601	EST975718
BC451050	NF108B04D
BG415233	RH122_22-
AL053849	RH122_12-
AI47672	Anopheles
BI269364	NF008A07F
BI269344	NF005H101

V4040408 734 bp mRNA EST 06-FEB-2000
 V4040408 Bombyx mori pheromone gland Shuko x Ryuhaku new1y-closed
 adult Bombyx mori cDNA clone pg-0946 T3, mRNA sequence.
 V4040408
 V4040408.1 GI:5908496
 EST.
 Domestic silkworm.
 Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Lepidoptera; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
 ; Bombycoidea; Bombycidae; Bombyx.
 1 (pass 1 to 734) Shimada, T., Okano, K. and Maeda, S.
 Mita, K., Moriyono, M.,
 Bombyx mori cDNA
 Unpublished (2000)
 Contact: Mita K
 Genome Research Group
 National Institute of Radiological Sciences
 Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
 Email: kmita@nirs.go.jp
 method:uni-directional, sequence direction:sequenced from T3 primer
 (5' -> 3')
 Project:Silkworm Genome Program in MAFF, and Research for the
 Future Program in JSPS, see 'Silkbase',
<http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.
 Location/Qualifiers

```

1..754
/organism="Bombyx mori"
/strain="Shuko x Ryuhaku"
/db_xref="taxon:7091"
/clone="pg-0946"
/clone_lib="Bombyx mori pheromone gland Shuko x Ryuhaku"
/newly_enclosed adult"

```

/sex="female"
 /tissue_type="pheromone gland"
 /dev_stage="newly-eclosed adult"
 BASE COUNT 173 a 200 c 160 g 201 t
 ORIGIN

Query Match 77.6%; Score 724.4; DB 10; Length 734;
 Best Local Similarity 99.2%; Pred. No. 3.7e-211;
 Matches 728; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 170 tccggtctgcgtccggtatcgcttaattcagcaacgaaagacgctgcaacca 229
 Db 1 TCCGGTCTGCTTCGGTATTCGCTTAATTCAGCAACGAAAGACGACCTGGCTAACCA 60
 QY 230 gcttcgacgactcttcacgattacacacacgctgcttactctgttgcagatataa 289
 Db 61 GCTTCGCGGAGCTCTTCACGATTCACGACATGCTCTTACCTGTGTGCAATATTA 120
 QY 290 aatccgaaacgcttctgacgctcactatctacgttcgacgcttcgagattcatt 349
 Db 121 AATCCGAACGCTTATGACAGCTCTTAACATTAACCTGCGAACTTTCGGATTCATT 180
 QY 350 ttgcagacccctctcctgcgtcgaatgggttcagacgacatcgttcgaatgaacca 409
 Db 181 TTGCAACATCTCTGCTGCAATGTTGAGTTCACACATACGTCGAAGTACCAACT 240
 QY 410 aagcggaatcgttagtaagcgcgcctcttctacacacacacgacgacgacgac 469
 Db 241 AGCGGAATCGGTACTAGCCGCCCTCTTTCATCTCACTACCAACGACGCAATTA 300
 QY 470 cccatcgttgagtaaaatttaacccaatttatacaataagtaacatcgcgttaata 529
 Db 301 CCCATCGTTGAGTCAAAATTAACCAATTTATCAATAGTCAATATATCCCTTATA 360
 QY 530 ttgtgcacatcgttgcacatcgtcgtacgctgacgacgacgacgacgacgacgac 589
 Db 361 TGTGTCATATGTCGCAATCATCTGCTACGTCGTCGTCGTCGTCGTCGTCGTCG 420
 QY 590 attgacatctcttctgacgacgacgacgacgacgacgacgacgacgacgacgac 649
 Db 421 ATTGACTCTTCTTGTGCGATTTGACCAACGACGTCATACACGCGCTTAAACAGTGC 480
 QY 650 gaccagtggtgttggttaagtttggatagacgttcgacgacgacgacgacgacgac 709
 Db 481 GACCAAGTGGTGGTAAAGTTGGGATTAGCATCTTCACAGCGCAATATGCTGGCTT 540
 QY 710 gctgacatcctgaatagcagacgacgacgacgacgacgacgacgacgacgac 769
 Db 541 GCTGGATCTTGAATATACGACGCTTTCATCTTCGACACTCTTTCGCAAAATCTT 600
 QY 770 ccccaacgctctgttttggcaataatacaacgacgacgacgacgacgacgacgac 829
 Db 601 CCCACACACTCTGTTTGGCAATATCAACGCGCTGTTACCAATGCAATCTCTCAT 660
 QY 830 ctgcgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 889
 Db 661 CTTCGCCCGCGCGTCGCGGACACGCAATATCCGATTAAGCAATGTGGAGACAT 720
 QY 890 tgcagtaactcttgc 903
 Db 721 TGCAGTACCTTTGC 734

RESULT 2
 AIO65168/c 628 bp mRNA EST 24-JUL-1998
 LOCUS AIO65168
 DEFINITION ESTRNA-8 Rat Lambda Zap II Library, StrataGene #36501 Rattus
 norvegicus cDNA clone pGEM-T/TRKA-8MF 3', mRNA sequence.
 ACCESSION AIO65168
 VERSION AIO65168.1 GI:3340575
 KEYWORDS EST.
 SOURCE Norway rat.

ORGANISM

Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 628)

AUTHORS

Bingdong/L. and Huang/B.R.

TITLE

Research on mechanism of p/53NMR induced apoptosis (Bingdong/L. and Huang/B.R.)

JOURNAL

Unpublished (1998)

COMMENT

Contact: Dongbing Lai
 National Laboratory of Medical Molecular Biology
 Chinese Academic of Medical Sciences & PUMC
 5 Dongdan Sanliao, Beijing 100005, P.R.China
 Tel: 86-10-65296406
 Email: huangbr@med.mimc.ac.cn
 Seq primer: M13 Forward Primer.

FEATURES

Location/Qualifiers

source

1..628

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="pGEM-T/TRKA-8MF"

/sex="Male"

/tissue_type="Brain"

/dev_stage="6 weeks"

/lab_host="XLI-Blue MRF"

/note="Vector: Lambda Zap II; Site_1: EcoR I; Site_2: Xho I"

BASE COUNT

173 a 138 c 166 g 148 t 3 others

Query Match

Best Local Similarity 98.7%; Score 605; DB 10; Length 628;
 Matches 608; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 192 gcttaattcagcaacgaggaagacgactgctacacgacgcttcgacgacttcacgatt 251
 Db 628 GTTATTTTACGNCNCNMAAAGACACTGGCTACCAAGGCTCCGACTCTTCACGATT 569
 QY 252 atcgactcaatctcttcccttccctgtgttcagataataaaacacgaaacgctatgacg 311
 Db 568 ATCGACTCAATGCTCTTACCTGTTGTCAGATATTAATAAATCCGAAACCGTTATGCA 509
 QY 312 ctctaacttaactctgcgaacgcttcgagatgcaatttgcagacgctctgcctgcga 371
 Db 508 CTCTACTATTACTCTGCAACTCTTTCGCGATTGCAATTTTCAGACCTCTCTGCTGCGA 449
 QY 372 tgcgtgagttcagacgacgacgacgacgacgacgacgacgacgacgacgacgac 431
 Db 448 TGGTTGGAGTTCCAGAGCATACGTGCAAAATGACCAATGACGGAATCGCTACTAAGCC 389
 QY 432 cgcctcttctcatcttctacacacacgagacgacgacgacgacgacgacgacgac 491
 Db 388 CGGCTCTTTTCATCTCTACACCAACGAGCATTAACCATTTGTTGAGTCAAAATTAAC 329
 QY 492 ccaatttctcaataagtcacatcatcgcgttaataatgttgcacatccgctgcgaatcat 551
 Db 328 CCAATTTTATCAATAGTCAATATGCGCTTAAATATGTTGCCATTCGTCGCAATCAT 269
 QY 552 gctcgaagtgtagcgcgacgacgacgacgacgacgacgacgacgacgacgacgacgac 611
 Db 268 GCTGCTAACGTGTACACGCAATCAAAATGTTGCTCGATTCACCTCTTCTTGGGCAATT 209
 QY 612 gaccacacagagctcctacagcgcgttcaacgacgacgacgacgacgacgacgacgac 671
 Db 208 GCACACACAGGCTCTATACAGCGGCTTAACAGTGTGACGACGACGATGGGTAAAGGT 149
 QY 672 ttggattagacatcgtcagacgacgacgacgacgacgacgacgacgacgacgacgac 731
 Db 148 TTGGATTACATCTGTCACGCGCATATGCTGGCTTGGCATCTTGAATTAAGCGCA 89

Tue Jan 22 15:12:22 2002

us-09-740-211-13_copy_11000_11933.rst

OY 732 cgccttgacatcttcgcacatcttctgacacatctcccaagctctgtttggca 791
 Db 88 cgccttgacatcttcgcacatcttctgacacatctcccaagctctgtttggca 29
 OY 792 latcaacgcagcgc 807
 Db 28 TATCAACAGCAGCGCC 13
 RESULT 3
 BC451667 589 bp mRNA EST 16-MAR-2001
 LOCUS NF092610D1J1081 Drought Medicago truncatula cDNA clone NF092610D1
 DEFINITION 5', mRNA sequence.
 ACCESSION BC451667
 VERSION BC451667.1 GI:13370461
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 1 (bases 1 to 589)
 Torrez-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula drought library
 Unpublished (2000)
 JOURNAL Contact: May GD
 COMMENT Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7391
 Fax: 580 221 7380
 Email: gdmay@noble.org
 Insert Length: 589 Std Error: 0.00
 Plate: 092 row: 6 column: 10
 Seq primer: TCACACAGCAACAGCTATGAC.
 Location/Qualifiers
 FEATURES
 source
 1. 589
 /organism="Medicago truncatula"
 /db_xref="taxon:3860"
 /clone="NF092610D1"
 /clone_lib="Drought"
 /tissue_type="plantlets"
 /dev_stage="pooled timepoints"
 /note="vector: Lambda Zap; Contains a mixture of entire
 plantlets harvested in a series of days post-watering
 timepoints."
 timepoints 143 c 143 g 165 t
 BASE COUNT 138 a 143 c 143 g 165 t
 ORIGIN
 Query Match 63.0%; Score 588; DB 11; Length 589;
 Best Local Similarity 100.0%; Pred. No. 2.4e-169; Indels 0; Gaps 0;
 Matches 588; Conservative 0; Mismatches 0;

OY 329 gaactgttcggatgcatcttgcacacatctctgcctgcagatggttgagttccagac 388
 Db 242 GAACTGTTCCGGATTGCGATTTCGACACTCTCTGCTGCGATGTTGAGTTCCAGAC 301
 OY 389 gataagtcgaatgac 448
 Db 302 GATAAGTCGAATGAC 361
 OY 449 ctaccac 508
 Db 362 CTACCAC 421
 OY 509 gtcaataatcatgacgcatatagttgcatccatccgttgcaatcatgctcagtgacac 568
 Db 422 GTCAATATCAGTCGGTTAAATATGTTGTCATCCGTCGATCATCTGTTAAGCTGACAC 481
 OY 569 gaattcaaaatgtctgtcgtgcatgactctctctgtgcatgacacacacacacacac 541
 Db 482 GCATCAAAATGTTGTCGATGACCTCTCTTGTGTCGATGACACACACACACACAC 676
 OY 629 taacagcgttcaacagtcgtagacacacacacacacacacacacacacacacacac 589
 Db 542 TACAGCGCTTAACTGCTGCTGACACAGGTGGTGGTGAAGCTTTGGG
 RESULT 4
 A1489217 656 bp mRNA EST 18-MAY-2001
 LOCUS EST247556 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
 DEFINITION A1489217, mRNA sequence.
 ACCESSION A1489217
 VERSION A1489217.1 GI:4384588
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 656)
 Alcald, J., Vredalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E.,
 Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman
 , C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley
 , S.D. and Giovannoni, J.
 Generation of ESTs from tomato carpel tissue
 Unpublished (1999)
 JOURNAL Contact: CUGI
 COMMENT Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 Location/Qualifiers
 FEATURES
 source
 1. 656
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLED1717"
 /clone_lib="tomato ovary, TAMU"
 /tissue_type="carpel"
 /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
 /note="vector: pBlueScript SK(-); Site 1: EcoRI; site 2:
 XhoI; cLED - tomato Carpel EST library. OligoDT-primed and
 directionally cloned cDNA in vector Lambda Zap II with 5'
 and 3' ends located at the EcoRI and XhoI sites,
 respectively."
 respectively 159 c 153 g 176 t 2 others
 BASE COUNT 166 a 159 c 153 g 176 t
 ORIGIN
 Query Match 61.0%; Score 569.6; DB 10; Length 656;
 Best Local Similarity 99.0%; Pred. No. 1.1e-163;

```

/db-xref="taxon:4081"
/clone="CT0814J1"
/clone_lib="tomato flower buds 3-8 mm, Cornell University"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/notes="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; supplier: Tanksley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, 7d496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

BASE COUNT      157 a      124 c      135 g      130 t
ORIGIN

Query Match
Best Local Similarity 56.9%; Score 531.4; DB 10; Length 546;
Matches 543; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      65 ggcgcagcgcttaactactaataatgaatccagacagcagtgctcagctcagttggg 124
      |||
Db      545 GCGCGACGCGTAATCTATTCTATTATTATTCAGACAGACAGTGGCTACGGCTCAGTTGGG 486
      |||

QY      125 ttgtgtctgtctggcgcgagctagcgccttaacgaattggcgatccggtttctgtctc 184
      |||
Db      485 TTGTGCTGTGCTGGCGCGCGATGACCCCTGTACGCATTGGTGATCCGGTTCTGCTTCC 426
      |||

QY      185 ggtattgcttaattcagcacaaaggaagagcagctgctcctaacccagctcgcgcgactctt 244
      |||
Db      425 GGTATTGCTTAAATTCAGCAACAGGAAGACACCTGGCTAACAGGCTCGCGACACTTT 366
      |||

QY      245 cagcattatcgcactcaatgctcttactctgtgtgcagataataaaatcccgaaacggt 304
      |||
Db      365 CAGCAATTATGACTCAATGCTCTTACCTGTGTGTGAGATATAAATAATCCGAAACCGTT 306
      |||

QY      305 atgcagcgtctaactatcttaacctgcgaactgtttcggatggcattttgcagactctcgt 364
      |||
Db      305 ATGCAGCGCTCTCAATCTATTACCTCGCAACTGTTGGGATTTGCCATTTTGCAGACCTCTCTG 246
      |||

QY      365 cctgcgagtggttggaagttccagacgaatacgtgaaatgacaaactagagcgagatcgtag 424
      |||
Db      245 CCTGCAGATGTTGAGTGTGACAGAGATAGCTCAAAAGTGAACAACTAGCGGGAATCGGTAG 186
      |||

QY      425 taagcgccgcctctttcatctcactcaactacacagacagagaaatlaacccatggttgagca 484
      |||
Db      185 TAAAGCGCGCGCTCTTTCATCTACCTACTACCAACAGCAGCAATTAACCCATCTGTGATCA 126
      |||

QY      485 aatttaccacaatttcttaataaagtcacatcactgcgcgttaaatatgtttgcacatcgtag 544
      |||
Db      125 AATTTCACCAATTTTATTTCATTAAGTCAATTCATGCGCTTAATATGTTGCCAATCCGTGG 66
      |||

QY      545 caatactgttgttaac-gtgtaccgcactcaaaagtgtgtcgtgatatgactctctt 603
      |||
Db      65 CAATCATGCTGCTTAACGCTGTGACCGGCATTCAAATATGTTGTGCGATTAAGTCTTCTTT 6
      |||

QY      604 gtggc 608
      |||
Db      5 GTGGC 1

```

E

Db 301 CCCCACAGCTCTTTTGGCAATATACCGCAGCGCTGTCACTGACATCTCTGCAT 360
 QY 830 ctggcccccgcgtcgcgagcactacggaatataatccgataagcgaatgtggagcact 360
 Db 361 CTTGCCCCCGCGCGCGCAGCAGTACGCGCAATATATCGCATTAAGCGAATGTTGGACACT 889
 QY 890 tggagtaactcttgacttagtaattctctcaagctggccc 928
 Db 421 TGCAGTACCTTGGCTTAGTATTTCTTCAAGCTTGGCC 459
 RESULT 9
 BP328036/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; primates; Catarrhini; Hominidae; Homo.
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Goldmann, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Costa, F.F.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.G.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20020669
 Contract: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&cl=QV3-BN0149-SeqPrimer:puc18-forward>)
 High quality sequence start: 47
 High quality sequence stop: 617.
 Location/Qualifiers
 1. 617
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BN0149"
 /date_stage="Adult"
 /note="Organ: breast,normal; Vector: puc18; Site_1: Smal;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 158 a 152 c 151 g 156 t

us-09-740-211-13_copy-11000-11933.rst

[illegible]

RESULT	10				19-DEC-2000
BFG34302		685 bp	mRNA	EST	
LOCUS	BFG34302	685 bp	Drought Medicago	truncatula	cdna
DEFINITION	NF074F11D1T1094				clone NF074F11D1
	5', mRNA sequence.				

ACCESSION	Bf634302
VERSION	Bf634302.1
KEYWORDS	GI:11898460
SOURCE	
ORGANISM	
	barrel medic.
	EST.
	Medicago truncatula
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Eukaryota; Viridiplantae; eudicotyledons; core eudicots;
	Eukaryota; Viridiplantae; eudicotyledons; core eudicots; Trifoliaceae

Medicago.
1 (bases 1 to 685)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Weller, J.W. and May, G.D.
Imman, J.T., Weller, J.W. and May, G.D.
the Samuel Roberts Noble Foundation

Flores, H.N., 2000. Expressed Sequence Tags from the *Medicago truncatula* drought library. Unpublished (2000)

Unpublished (2003)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
P.O. Box 227
Ardmore, OK 73402, USA

Insert Length:	685	Seq Len:	685
Plate:	074	row:	F
Seq primer:	TCACACAGGAAACAGCTATGAC.	column:	11
Location/Qualifiers			

Seq primer: TCAACACAGT
Location/Qualifiers
1..685
/organism="Medicago truncatula"
3880"

```

/organism="taxon:3880"
/db_xref="NF074F11.DT"
/clone_lib="Drought"
/clone_lib="Drought"
/clone_lib="Drought"

```

```
/clone_id="Plantlets"  
/tissue_type="Pooled timepoints"  
/dev_stage="Lambda Zap; Contains a mixture of entire  
/note="Vector: Lambda Zap; Contains a series of days-post-watering
```

/note=vector; i = 1
plantlets harvested in a series of days f--
timepoints. "

	163 a	164 c	169 g	189 t
--	-------	-------	-------	-------

UNT	Timepoints:
163 a	169 g
164 c	189 t

OY	1	aaagcgcaaggaag	1
Db	260	AAACGCGCAGGAGTTGTATGCGCGACCTCCGACCCCGCTTTCACGAAAGTCAATGAA	120
OY	61	aaagcgcgcaagctaacctatactaattcagacagaaagatggcctacgctcaagtt	379
Db	320	AAAGGCCCGCAGCGTACTATTACTAATTATTCAGACACAGAGTGGCTACGGCTCAGTT	180
OY	121	tgggtgttgctgtcttcgcygcgagatagacgcctgtacagcatttgtgtgcggtttcgc	439
Db	380	TGGGTGTGCTGTGCTGTGGGCGCGATGAGCCCTGTACGCATTGGTGAATCGGTTTTCG	240
OY	181	tccgcgtatccgttatatccagcaacaacggaaagacactgctaaacagagctcgcgac	499
Db	440	TTCGCGTATTGCGTTAATTCACACACACGGAAGAGACACTGGCTAAACACGGCTGCGCGAC	300
OY	241	tcttaacagattacgcactcaagtcctcttaccctgttgtcaagtatataaanaatccgaac	559
Db	500	TCTTTCACGATTATCGACTCAATGCTTACTCGTGTGGCAGATATMAAAATCCCGAAAC	360
OY	301	cgttatagcagagctcaactatcactcgcgaactgtttcggatctgttcattttgcagacctc	619
Db	560	CGTTATCGAGGCTCTAATCTATTACCTGCGCAATGTTTGGATGGATTTGTCAGACCTTC	420
OY	361	tctgcgtcgatagtggttgagatgccagacgatactgcgaagtgaaccaactaagcggaaatcg	679
Db	620	TCMGCTGCGATGCTTGAGTTCACAGACGATTACGTGGAAGGACCAACTAGCGGGAAATCG	779
OY	421	9 421	
Db	680	G 680	

RESULT 11	EST	18-JUL-2001
BI269141	688 bp mRNA	
LOCUS	MF003405.1	tridactylated Medicago truncatula cDNA clone
DEFINITION	MF003405.1	5' mRNA sequence.

ACCESSION
 B1269141
 VERSION
 B1269141.1
 KEYWORDS
 EST,
 barrel medic.
 SOURCE
 Medicisio truncatula
 ORGANISM
 Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eudicots; core eudicots; rosids; Malvales; Malvaceae;

Eukaryota; Eucarya; Eukaryotophyta; Magnoliophyta; Spermatophyta; Fabaceae; Fabaceae; Papilionoideae; Rosidae; eurosids I; Fabales;

Medicago:
1 (bases 1 to 688)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell,
W.J., Weller, J.W. and May, G.D.
from the Samuel Roberts Noble Foundation
Flores, H.R., Imman, J.T.,

TITLE
Expressed Sequence Tags from
Medicago truncatula irradiated library
reestablished (2001)

JOURNAL
COMMENT
Unpublished
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
P.O. Box 73402, USA

THE Ssm Noble Parkway, Atlanta, GA 30328
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert length: 688 Sld Error: 0.00
Plate: 003 row: A column: 05
Seq primer: TCACACAGCAACAGCTATGAC.
Location/Qualifiers

```
FEATURES
SOURCE
1. .688
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF003A051R"
```

/clone.lib="Irradiated"
/issue.type="seedlings"
/dev.stage="seedling"
/note="vector: Lambda Zap: Seedlings were exposed either gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated at 6, 12, 24 and harvested 24 hours post treatment. cDNA samples were polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was prepared from ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. from the recombinant Uni-Zap XR vector using Exsist (Stratagene). Excised plasmids were plated using SOLR cells."

BASE COUNT 140 a 193 c 155 g 193 t 7 others
ORIGIN

Query Match
Best Local Similarity 41.5%; Score 387.2; DB 11; Length 688;
Matches 411; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Db 515 atcagccgttaataatggtccatccgtgcaatcagc-tgctacgtgtgacgcat 573
1 ATCATGCCGTTATATGTTGTCATCCGTCGCAATCAATGCTGTACGTGACCGCAT 60
Qy 574 caaatgtgtctgcatgactctcttctgtgcatgacacacacagagcgatcatt 60
61 CAAATGTTGTCGATGATGACTCTTCTTGTGGCATTTGACATTTAGCATGTACAGC 120
Qy 634 cggatcaagatcgctgacagcagtgagtgagtgagtgagtgagtgagtgagtgag 693
121 CGGCTTAACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy 694 gcaatagc-tgagctgctgcatccttgaatagcagcagccttgccttcagcagc 693
181 GCGATAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 180
Qy 753 cttctcgcaactcctcccccagcctctgttctgcaatatacagcagcagccttgc 752
241 CTTTCTCGACACTCTCCCTCCAGACTCTGTTTGGCATATCAACGCGCTGTAC 240
Qy 813 catgcaactcctgcatcttgcctccgagctgagcagcagcagcagcagcagcagc 300
301 CATGGCATCTCTGATCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 872
Qy 873 gcaatgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360
361 CGATGTTGCGACACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 928
Db

RESULT 12
LOCUS AV404200 775 bp mRNA EST
DEFINITION AV404200 Bombyx mori pheromone gland Shuko x Ryuhaku newly-eclosed
ACCESSION AV404200
VERSION AV404200.1 GI:6908288
KEYWORDS EST.
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
REFERENCE 1 (bases 1 to 775)
AUTHORS Mita, K., Morimoto, M., Shimada, T., Okano, K. and Maeda, S.
JOURNAL Unpublished (2000)
COMMENT Contact: Mita K
Genome Research Group

National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmil@nirs.go.jp
method: uni-directional, sequence direction: sequenced from T3 primer
(5' -> 3')
Project: Silkworm Genome Program in MAF, and Research for the
Future Program in JSPS, see "SilkBase",
http://www.db.a.u-tokyo.ac.jp/silkbase/, for whole ESTdb.
Location/Qualifiers
1. 775
/organism="Bombyx mori"
/strain="Shuko x Ryuhaku"
/db_xref="taxon:7091"
/clone="pg-0693"
/clone.lib="Bombyx mori pheromone gland Shuko x Ryuhaku"
/sex="female"
/issue.type="pheromone gland"
/dev.stage="newly-eclosed adult"

BASE COUNT 161 a 211 c 185 g 218 t
ORIGIN

Query Match
Best Local Similarity 40.7%; Score 380.2; DB 10; Length 775;
Matches 382; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 544 gcaatcagcctgcatcagcagcagcagcagcagcagcagcagcagcagcagcagc 603
1 GCAATCAGCTGCTTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Qy 604 gttgcatgac 603
61 GTGGCATTTGAC 120
Qy 664 ggttaggttggatgatacagcagcagcagcagcagcagcagcagcagcagcagcagc 120
121 GGTAGGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Qy 724 atagcagcagccttgcagcagcagcagcagcagcagcagcagcagcagcagcagc 180
181 ATAGCGACGCTTGTGATCTTCCGACTCTTCTCGACAACTCTCCCGACAGCTGT 240
Qy 784 ttgcaatatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
241 TTGCAATATCAACCGACGCGCTGTACCATGCAATCTGCAATCTTGGCCCGGCGT 300
Qy 844 cgcgacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
301 CCGGCACTACGCAATATATCCGATTAAGCAATGTTGCGACACTTGCAGTACCTTTGC 903
Qy 904 cttagatcttcccaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 360
361 CTTAGTATTTCTTCAAGCTTTTGC 385
Db

RESULT 13
LOCUS BG312470 373 bp mRNA EST
DEFINITION WHE2408_F10_L120S Wheat 3-6 DAP seed cDNA library Triticum aestivum
ACCESSION BG312470
VERSION BG312470.1 GI:13114273
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae
REFERENCE 1 (bases 1 to 373)
AUTHORS Akhunov, E., Anderson, O.D., Chao, S., Chiu, A., Choi, D.W., Close, T.J.,
Dvorak, J., Fenton, R.D., Gustafson, P., Han, P.S., Hsia, C.C., Kang, Y.,
Kianian, F., Lazo, G.R., Miller, R., Otto, C., Nguyen H.T., Rausch, C.J.,

TITLE , Seaton, C.L., Simons, K., Tong, J.C. and Zhang, D.
The structure and function of the expressed portion of the wheat
genomes - 3-6 DAP seed cDNA library
Unpublished (2001)
JOURNAL Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@wp.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stralagene SK primer.

FEATURES
source
1. 373
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone_lib="WHE2408.F10.L20"
/clone_lib="Wheat 3-6 DAP seed cDNA library"
/clone_type="Young seed"
/dev_stage="Adult plant"
/lab_host="E. coli SGR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Total RNA was prepared from
3-6 DAP seed (Gustafson), equal portions of RNA were
pooled from these four DAPs, polyA was purified, a cDNA
library was made, and the cDNA clones were in vivo
excised to give phagescript phagemids in the phi Close lab
at the University of California, Riverside (Akhunov, Chin
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."

BASE COUNT
87 a 93 c 93 g 100 t
ORIGIN

Query Match 39.9%; Score 373; DB 11; Length 373;
Best Local Similarity 100.0%; Pred. NO. 1.8e-103;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

36 cccgcttcacgaagtcagtgtataaagccgcagcgaactattactaatgaatcag 95
|||||
1 CCGCCTTTCAGAGTCAATGTGTAAGCGCGCAGCTTACTATTGATGATTCAG 60
96 gacacagatgctacagctcagttggtgtgtgtgtgtgtgtgtgtgtgtgtgt 155
|||||
61 GACAGACAGTGGCTACGCGTCACTTGGGTTGTGCTGTTCCGGCGCGATGACGCTG 120
156 taccgattgtgacccggtctgtcttcggtatctgcttaattcagacaaaggaaga 215
|||||
121 TACGCAATTTGGTGATCGGTTGCTCCGATTCGCTTAATTCAGCAACGGAAGA 180
216 gaactgctaaccaagctgcgcgactcttcagattacgaactcaatgcttaccgtt 275
|||||
181 GCATGCGCTTACACAGGCTCGCGACTCTTACGATTAATGACATCATGCTTACTCTGT 240
276 gtgcagataataaataccgaacacgttatgcagcgtctaatattactattcgaactgt 335
|||||
241 GTGCGATTAATAAATCCGAACCGTTATGCGGCTCACTATTAATTAATCCGGAACCTGT 300
336 ttgcagattgcaatttgcagacacctgtgcctgcagatgattgagttccagacagatcgt 395
|||||
301 TTGCGGATTTGCAATTTGACAGACCTCTCTGCTCGATGATTGGAGTTCAGACGATACGT 360
396 cgaagtgcacaaac 408
|||||
361 CGAGTGACCAAC 373

RESULT 14
AA841579

LOCUS AA841579 498 bp mRNA EST 02-MAR-1998
DEFINITION MB3D6AA9E1173 Brugia malayi day 6 post-infection third stage larvae
SAM96MMLW-BML3d6 Brugia malayi cDNA clone 3D6AA9E11 5', mRNA
sequence.

ACCESSION AA841579
VERSION AA841579.1 GI:2922915
KEYWORDS EST.
SOURCE Brugia malayi.
ORGANISM Brugia malayi.
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
Onchocercidae; Brugia.

REFERENCE 1 (bases 1 to 498)
Blaxter, M.L., Waterfall, M., Daub, J., Lizotte-Waniewski, M., Baron, L.
and Jones, S.J.
Genes expressed in day six post-infection, third stage larvae of
Brugia malayi

JOURNAL Unpublished (1997)
CONTACT: Blaxter ML
INSTITUTE OF CELL, ANIMAL AND POPULATION BIOLOGY
UNIVERSITY OF EDINBURGH
ASTMOTH LABS, KING'S BUILDINGS, WEST MAINS ROAD, EDINBURGH, EH9
3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The ABI trace of this sequence can be viewed at
<http://www.sanger.ac.uk/Brugia/3D6/MB3D6AA9E1173.html>
Seq primer: T3.

FEATURES
source
1. 498
/organism="Brugia malayi"
/strain="T3S Labs"
/db_xref="taxon:6279"
/clone_lib="3D6AA9E11"
/clone_lib="Brugia malayi day 6 post-infection third stage
larvae SAM96MMLW-BML3d6"
/sex="mixed"
/dev_stage="third stage larvae, six days after infection"
/lab_host="E. coli XL1-Blue"
/note="Vector: LambdaZap XR; Site 1: Eco R I
(5' end); Site 2: Xho I (3' end); Brugia malayi is a
lymphatic filarial nematode parasite of humans. mRNA was
prepared from third stage larvae of Brugia malayi isolated
from the peritoneal cavity of jirds six days after
infection. The mRNA was converted to double stranded cDNA
using reverse transcriptase and oligo(dT) followed by
RNase H and DNaPol I. The library had 2 x 10⁶ independent
recombinants and average insert size was 900 base pairs.
The library is available from The Filarial Genome Project
Resource Center: contact Dr. S.A. Williams, Clark Science
Center, Smith College, Northampton, MA 01063 USA phone +1
413 585 3826 fax +1 413 585 3786 email genome@smith.edu"

BASE COUNT
102 a 139 c 123 g 134 t
ORIGIN

Query Match 39.8%; Score 371.8; DB 10; Length 498;
Best Local Similarity 99.5%; Pred. NO. 4.8e-103;
Matches 373; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

96 gacacagatgctacagctcagttggtgtgtgtgtgtgtgtgtgtgtgtgtgt 155
|||||
124 GCGTGACAGTGGTACGCGCTCAGTTGGGTTGCTGTCGCGCGCGATGACGCTG 183
156 taccgattgtatccggttctgtcttcggtatcgtttatcagacaaaggaaga 215
|||||
184 TACGCAATTTGGTGATCCGGTTGCTTCGCGATTCGCTTAATTCAGCAACGGAAGA 243
216 gaactgctaaccaagctgcgcgactcttcagattacgaactcaatgcttaccgtt 275
|||||
244 GCATGCGCTTACACAGGCTCGCGACTCTTACGATTAATGACATCATGCTTACTCTGT 303

RESULT	15
LOCUS	AV749428
DEFINITION	AV749428 399 bp mRNA EST
ACCESSION	AV749428 NPC Homo sapiens cDNA clone NPDBD10 5', mRNA sequence.
VERSION	AV749428.1 GI:10907276
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 399)
AUTHORS	Song,H., Peng,Y., Gu,Y., Yang,X., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu, ,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q., , Han,Z., Chen,Z., Hu,R., and Chen,J. Homo sapiens NPC library cDNA clones Unpublished (2000) Contact: Qinghua Zhang Shanghai Institute of Endocrinology, Rui-Jin Hospital 157 Rui-Jin II Road, Shanghai 200025, P. R. China Tel.: 86-21-64370045(ex. 663332) Fax: 86-21-64743206 Email: mbsl@ems.stn.sh.cn This clone is available at Shanghai Hematology Institute in Changhai. Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong. Location/Qualifiers 1..399
FEATURES	
source	

Query Match	38.5%;	Score 360;	DB 10;	Length 399;
Best Local Similarity	96.3%;	Pred. No. 1.8e-99;		
Matches 388; Conservative	0.0;	Mismatches 13		

[illegible]

```
Search completed: January 17, 2002, 14:58:51
Job time: 8652 sec
```

Tue Jan 22 15:12:22 2002

us-09-740-211-13_copy_11000_11933.rst

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1000	100.0	11933	6	AR138377	AR138377 Sequence
2	1000	100.0	11933	6	AR146887	AR146887 Sequence
3	725.2	72.5	4399	6	AR138378	AR138378 Sequence
4	725.2	72.5	4399	6	AR146888	AR146888 Sequence
5	577.4	57.7	5035	6	AR034084	AR034084 Sequence
6	577.4	57.7	5035	6	AR071306	AR071306 Sequence
7	577.4	57.7	5035	6	AR112722	AR112722 Sequence
8	575	57.5	8241	6	AR05328	AR05328 Synthetic F
9	575	57.5	8241	6	AR07042	AR07042 Artificial
10	575	57.5	8831	6	E00527	E00527 Complete cd
11	575	57.5	8967	6	AX052730	AX052730 Sequence
12	575	57.5	8967	6	171409	171409 Sequence
13	575	57.5	8967	9	HSRVITIR	X01179 Human mRNA
14	575	57.5	9229	6	H05VITIC	M14113 Human coagu
15	574	57.4	7272	6	I05404	I05404 Sequence
16	573.4	57.3	9009	6	AR003710	AR003710 Sequence
17	573.4	57.3	9009	6	AR029065	AR029065 Sequence
18	573.4	57.3	9009	6	AR126884	AR126884 Sequence
19	573.4	57.3	9009	6	I31901	I131901 Sequence
20	573.4	57.3	9009	6	I63424	I63424 Sequence
21	573.4	57.3	9009	9	H05VITIR	K01740 Human coagu
22	572.4	57.2	7056	6	I27063	I27063 Sequence
23	570.2	57.0	7440	6	I08345	I08345 Sequence
24	567.8	56.8	9354	6	AR003585	AR003585 Sequence
25	562.2	56.2	7440	6	I08457	I08457 Sequence
26	517	51.7	4278	6	I08644	I08644 Sequence
27	517	51.7	4281	6	I08643	I08643 Sequence
28	517	51.7	4581	6	I08642	I08642 Sequence
29	517	51.7	4551	6	I08641	I08641 Sequence
30	517	51.7	4670	6	AR110040	AR110040 Sequence
31	517	51.7	6999	6	I77105	I77105 Sequence
32	501.4	50.1	19343	6	I02047	I02047 Sequence
33	444.2	44.4	7032	4	AF016234	AF016234 Canis fam
34	443	44.3	7493	6	AR003712	AR003712 Sequence
35	443	44.3	7493	6	AR029067	AR029067 Sequence
36	443	44.3	7493	6	AR126886	AR126886 Sequence
37	443	44.3	7493	6	I63427	I63427 Sequence
38	443	44.3	7493	10	MSACVITII	L05573 Mus domest
39	442.6	44.3	7145	4	AF049489	AF049489 Canis fam
40	414.8	41.5	6539	4	SSM49517	Y49517 Sus scrofa
41	413.8	41.4	4334	6	AR029058	AR029058 Sequence
42	413.8	41.4	4334	6	AR126917	AR126917 Sequence
43	413.8	41.4	6402	6	AR029097	AR029097 Sequence
44	413.8	41.4	6402	6	AR126916	AR126916 Sequence
45	371.2	37.2	553	4	AF180523	AF180523 Ovis arie

ALIGNMENTS

PAT 16-JUN-2001

RESULT	1	DNA
ARI38377	11933 bp	US 6200560
LOCUS	Sequence 13	from patent

DEFINITION	ARI38377	GI:14480722
ACCESSION	ARI38377	
VERSION	ARI38377.1	

SOURCE	ORGANISM	TITLE	REFERENCE	AUTHORS	JOURNAL
Unknown.	Unclassified.	1 (phases 1 to 11933)	Canto, L.B., Colosi, P.C., and Qian, X.	Adeno-associated virus vectors for expression of factor VIII by target cells	Patent: US 6200560-A 13 13-MAR-2001;

Source	1. .11933	/organism="unknown"
BASE COUNT	3258 a	2818 c 2717 g 3140 t
ORIGIN		

QY	361	tgcgctgcttgaattacatgacatgacatccatcttcttcttcttccacagatgat	420
Db	351	TCCTAAATACCACTACTGACTGACATCCACTTTTCTTTCTCCACAGATATCA-	409
QY	421	tcacacatgcaaaatagagctctcacctgctcttcttcttcttcttgcattgcatctt	480
Db	410	TCACACATGCAAAATAGAGCTCTCCACTGCTCTTTCTGTGCTTTTGCGATTGCTTT	469
QY	481	agtgccacagaatatactacctgggtgctgagtggaactgtcatggagcatatgcaagt	540
Db	470	AGTGCCACCAAGAATACTACTCTGGGTGCAATGACACTGTCAATGGGACATATAGCAAACT	529
QY	541	gactcgtgtgacgtgcctgtgagcagaatattccctctagagtgcacaaacttttcca	600
Db	530	GACTCTGGTGAAGCTGCTCTGTGACGACAGATTTCCTCTTACAGTGCCAAAATCTTTTCCA	589
QY	601	ttcaacacatcagtcgtgtacaaaagaactctgtttgtagaattcacggatcacctttc	649
Db	590	TTCAACACCTCACTCTGTATACAAAAGACTCTGTTGTATGAAATTCACGATCACCTTTTC	649
QY	661	aactgtcttagcagaagcgcaccccttgatgtgtctgtctaggtccctacatccagctgag	720
Db	650	AACCTTCCTTAGCCCAAGCCACCCCTGGATGGTGTCTAAGTCTTACCATCCAGCCTTAG	709
QY	721	gttatgatcagtggtgcatcttaactaaagaacatggtctccatccctgtgagttcat	780
Db	710	GTATATGATACAGTGCTCATTAACCTTAAGAACATGCTCCATCTCTGCTACTTTCAT	769
QY	781	gctgtgtgtatccctactggaagctctctgagggagctgaatagatgacagaccagt	840
Db	770	GCTGTGTGTCTCTCTACTGTGAAGCTTTCTGAGGAGCTGAATATATGATGTCGACCACT	829
QY	841	caaggaggaaagaatgataaagcttccctggttgaagccatatacatatgcttgcag	900
Db	830	CAAGGAGAAAGAAGATGATTAAGTCTTCTCTGTGTGAAGCCATATATGATGTCGACAG	889
QY	901	gtccctaaagagaatggtccaatgagcctctgagccactgtgcttactactatcatctt	960
Db	890	GTCCTGAAGAAGATGTCTCAATGGCTGTGACCCACTGTGCTTACTACTCATATCTT	949
QY	961	tctcatgtgacctggttaaaagacttgaaattcaaggcccca	1000
Db	950	TCTCATGTGACCTGTGTAAGACTTGATTTACAGGCTCTA	989
RESULT 4			
LOCUS	ARI46888	4999 bp	DNA
DEFINITION	Sequence 14 from patent US 6221349.		PAT
ACCESSION	ARI46888		
VERSION	ARI46888.1	GI:15110691	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 4999)		
AUTHORS	Conco,L.B., Colosi,P.C. and Qian,X.		
TITLE	Adeno-associated vectors for expression of factor VIII by target cells		
JOURNAL	Patent: US 6221349-A 14 24-Apr-2001;		
FEATURES	Location/Qualifiers		
SOURCE	1..4999		
BASE COUNT	1375 a 1151 c 1155 g 1318 t		
ORIGIN			
QY	1	caagctagcgtctgctcgtctacatgagccgcggcgaagcccggtgctgagcagacc	60
		72.5%; Score 725.2; DB 6; Length 4999;	
		Query Match 85.9%; Pred. No. 6,7e-210;	
		Best Local Similarity 0; Mismatches 118; Gaps 4;	
		Matches 859; Conservative	

[illegible]

VERSION	KEYWORDS	AR034084.1	GI:5949689
SOURCE	ORGANISM	Unknown.	Unknown.
REFERENCE	AUTHORS	1 (bases 1 to 5035)	Unclassified.
TITLE	JOURNAL	Voorberg,J.J.	Hybrid proteins with modified activity
FEATURES	source	Patent: US 5869292-A 1 09-FEB-1999.	Location/Qualifiers
		1..5035	/organism="unknown"
BASE COUNT	1484 a	1127 c	1110 g
ORIGIN			1314 t
Query Match			
Best local similarity	57.7%	Score 577.4;	DB 6; Length 5035;
Matches 578; Conservative	99.8%;	Pred. No. 8.8e-165;	
	0; Mismatches	1; Indels	0; Gaps
QY	422	ccacatgcgaataagagctgcacactgcttcttcctgctgcttttgagatcagatcgttta	48
Db	30	ccaccatgagaaatagagcttccacactgcttcttcctgctgcttttgagatcagatcgttta	48
QY	482	gtgcaccagaagaatactactcctctgggtgcagtcggaacgtcatatggactatagaagtg	89
Db	90	gtgcaccaccaaaatgatactcctctgggtgcagtcggaacgtcatatggactatagaagtg	54
QY	542	atctctgtagctgcctctgtagacggaagatcttcctcctagtcagtcgaataatcttccat	601
Db	150	atctctgtagctgcctctgtagacggaagatcttcctcctagtcagtcgaataatcttccat	601
QY	602	tcaaacctccagtcgtgcatacaaaagactcgtcttgtagaatcagatcaacttttca	205
Db	210	tcaaacctccagtcgtgcatacaaaagactcgtcttgtagaatcagatcaacttttca	205
QY	662	acatgcgtaagcaagagccaccccttgatggtctgctgagtcctacccaacagctgag	721
Db	270	acatgcgtaagcaagagccaccccttgatggtctgctgagtcctacccaacagctgag	721
QY	722	ttatgataaagtggtcatctaaactaaagaatgcttccatccctgtaagcttcacatg	329
Db	330	ttatgataaagtggtcatctaaactaaagaatgcttccatccctgtaagcttcacatg	329
QY	782	ctgtctgtagtactcctacgtgaagactcttgaggagagctgaatatgatagcagaccagtc	389
Db	390	ctgtctgtagtactcctacgtgaagactcttgaggagagctgaatatgatagcagaccagtc	389
QY	842	aaagggagaaagaatgataaagtcctccctcctggttggaagcacaataatgctctggc	449
Db	450	aaagggagaaagaatgataaagtcctccctcctggttggaagcacaataatgctctggc	449
QY	902	tctctgaagaagaatggtccaaatggcctcctgaccacatgctgaccttaccataatctt	509
Db	510	tctctgaagaagaatggtccaaatggcctcctgaccacatgctgaccttaccataatctt	509
QY	962	ctcatctgagactctgtaaaagacttgaatcaggacctca 1000	
Db	570	ctcatctgagactctgtaaaagacttgaatcaggacctca 1000	
RESULT	6	AR071306	5035 bp DNA
LOCUS	DEFINITION	Sequence 1 from patent US 5910481.	PAT
VERSION	ACCESSION	AR071306	18-FEB-2000
KEYWORDS	KEYWORDS	AR071306.1	GI:7222194
SOURCE	SOURCE	Unknown.	
ORGANISM	ORGANISM	Unknown.	
REFERENCE	REFERENCE	Unclassified.	
AUTHORS	AUTHORS	1 (bases 1 to 5035)	
		Voorberg,J.J.	

```

CDS
    /cdoos start=1
    /transl_table=1
    /product="Factor VIIr"
    /protein_id="CA00448.1"
    /db_xref="GI:345062"
    /translation="MQLSTGCEFLICRRCSEATRRYYLGAVELSDMDSGLCEP
VDAEPFRRVRRSEFNTSYKPTFYETHTDENLIRKPPPMGLIGFTLOAEKTV
VVIILKKNASHPRSLAIVGYSTWKNSEGA EYDDOTSOREKEDDKVFGGSHTYMVI

```

Query Match	57.58;	Score 575;	DB 6;	Length 8241
Best Local Similarity	100.0%;	Pred. No.	5.1e-164;	
Matches	575;	Conservative	0	

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523</
--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------

Qy	846	gagagaagaatgatataaacgtctcccttgtaggaaccatatatgtctgcaggctc	905
Dd	426	GGAGAAAGAATGATTAAGCTTCCTCCGTGGAMACCATACATATGCTGCAGGTTCT	485
Oy	906	gaaagaagtgtccaattggcctcttgcaccactgttccttaacctatactttcca	965
Dd	486	GAAAGAGATTGTGCCAATWGCGCTTGACCACTGAGCCTTACTACTCATATCTTTCTCA	545
Oy	966	tgtgacctgttaaagaacttgaatcgacctca 1000	
Dd	546	TGTGGACCTGCTMAAGACTTGAATTCAGGCGCTCA 580	
RESULT	9		
LOCUS	A07042		
DEFINITION	Artificial mRNA for factor VIII.	PAT	24-AUG-1993
ACCESSION	A07042		
VERSION	A07042.1 GI:412967		
KEYWORDS	factor VIII.		
SOURCE	synthetic construct.		
ORGANISM	artificial construct.		
REFERENCE	artificial sequence.		
AUTHORS	Van Ooyen,A.J.J., Andreoli,P.M., Van Mourik,J.A. and Pannekoek,H.		
TITLE	I (bases 1 to 8241)		
JOURNAL	Method for the preparation of proteins with factor VIII activity by microbial host cells: expression vectors, host cells, antibodies GIST-BROCADES N.V. 1000, N.Y.		
FEATURES			

1. 8241
 /organism="synthetic Construct"
 /db_xref="taxon:23650"
 7. 7062
 /codon_start=1
 /transl_table=11
 /product="factor VIII"
 /protein_id="CAA00607.1"
 /db_xref="GI:412968"
 /translation="MQLSTLSTPCFSAIRRYVYGVVGLSMDWSQDGLP
 VLTIKNMASHSPFTNVSVKLTFLFELTFRHSDINLNIAKPRPMYGLGPTIAEYDT
 KENGPAADSLCTVSYSHDLYVDNLSDKSECAEEDTQSRREEDKTPGGSHIAEYD
 AADPEKSHWSTGKSNLMQDDASASARMPKRNITVNGYVNSLAEKTLGKTLFLF
 IAGCTIPFVHSLFLEHGTFLVNRNRSALKEISPTPLTQKLTMLDGLGFLFCHISSH
 OHGMFAVKKVDSCEEEQMLKKNNEAEEDDLTDSMDVARDDDNSDSFQIS
 TETRTTRALVHSHGSLGLGLGEVGDITLITFKKQSLRNPNPIIGKKYKVFPMY
 RLPKOVNKLQOELPEIETPKRYKATVDPGKTSQPCPLCTVYVTSFVPMNRDGLASL
 IDCLLTCKRESDVNGDQMSDRNVLTVSYDENKSMVLENTLQRELPYAGVAGLSD
 PEQASNMHSINQVYRDSLOLSCLEHVAITVLSIAQDPLFVSEFGSGTKHHYV
 YEDTLTLPFGEFGEVPMENPCLTGLGHNSDFPNRMATLAVSSCDKNTGDEYD
 SEDTSLAVLSEKNNALIEPRSSNSHSTROKQONATTPENDIETDPMFAPHPM
 KPLHNSVSDMLTLMKQSLPHTLRLNEKGLTAAATELKLDFKVSSTNNLSTISDILA
 RPOHNSGSLVPMSPHYVQSOLDTLTGFKKSSPLTEGSGTSLSEENKSNLTLESGLM
 NSQSSGSGKLVSESGREKFKKRRAGPALLTKNALFETKSLTALTKNTSNKSTNR
 KTHIDGSLITLHNSPSYMONLTLESOTFERKCTPLIHRLMLDKNATLAKNTNTSNKSTNR
 SSKMMEQVSEKGPILPPKQNPMDSPFKMLTLPESARWTOFSPKGNLSLNSQSGSPK
 OLTVSLPEVSYSEKGLPEKQNVVGVGKEFLQGLKREVPKRNGLNSLNSQSGSPK
 LTNHOKKIQEELERKLTIDENVLVQITLVGTGKNPKNLLPLSTKQNVGLNDLHEN
 YAVLDLPFSDSLNTRKLTIAFHKSGKEENLEGLGKNTQKMLTQKATRTIISPMP
 SOGNTPQSLKALQFPLPLETELKRIYVDTSTQSNKMKHLTSTLTQIDYVE
 KEGATLOSLDKLTJRSHPQENSRPLIAVVSPTSIRIYLTVALPDNSHSLP
 EASYTKKDSQVSESHFLQGGKKNLSIALTLTLEMGDQREVSGLGTATNSYVKKV
 KNNENRPGKPLVLTATRESSATPRKLDPLAMNHGTOPIPEENKSOEKSPERTA
 ITRTTLQSDQDEIYDQITISVKKKDDPIYEDNSQAGRTERLCKSTNHYIIVARE
 ITRTTLQSDQDEIYDQITISVKKKDDPIYEDNSQAGRTERLCKSTNHYIIVARE
 EVDENIIVTFEPNARQSPSTSLVSEYDQKQGEKPRKNRPNKTKYTKAIVHHM
 APTKDEDECKAMVPSYVDLDEKDVHSGILGPLVLCQHPNTLNPAGQVYVTEHFLFT

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1	(Bases 1 to 8967)	Amniota; Eutheria; Primates; Carnivora; Vertebrata; Euteleostomi; Saenko, E. L. and Strickland, D. K.	Methods of reducing factor viii clearance and compositions therefore	WO 0071714-A 1 30-nov-2000; The American National Red Cross (US)	Location/Qualifiers I. . 8967

```

/db_xref="taxon:9606"
110. 166
110. 7165
/note="unnamed protein product
/codon_start=1
/protein_id="CAC21811.1"
/db_xref="GI:12226921"
/translucio="two"

```

VITITLNNASHDPLVAVGVSYNKTLVEFDHJFNFKARPMWMLDPTIOQAVYDZ
 KNGMNAIDPLCTYLSTLSDVLDKJDNKSLIGALLTCNROBSGAKETROTJHFKITFLE
 IAGDFGKPMHSEITNANIMORDBASARPMKNGALTYCNVGSLEKROTJHFKITFLE
 IOMGTGTPVHISITPLEGHTJLVNHRQASITLSEITLTAQNTLMDQGLFCHISYMH
 OAKHGEAVATKSDPEEPOJRMKNNEADVDYDOLDEBEMVORFEDDNSSPTIORS
 VAKHPTKRWYHIAAEEDMDVAPLADPDRSKSOYLNNPDRIGRKKKPRMAY
 TDBETKTRJALHSEHISLLEPILVGVGVTLLITKRNASRYNINPGITDVAJLSR
 PLRGVNHLDPTLJLGEJFKKMYVETJEDPTSDRCPLRYTSSVNMVEDLASGL
 IGLPILCKESYVORCONIMSKRNVITLSEVFNDRSNVLTENIOFETPNVGOVLED
 PEFQASINMHSINGXVPSNOLASCLHEVAWITLSTGAOTDITLFSBFGYFKFKW
 YEDJLITLPEFSGEPGPNBNSNGMLTJGCHNSPFRNGMALTATVSCOKNGTYEED
 SYEDISAVYLLKMLNITERSSNORRSTROKFOFANTIPENDLEKTDMPFAHTPM
 PIKONVSSDILKMLNATERSSNORRSTROKFOFANTIPENDLEKTDMPFAHTPM
 RPLHSGGMPVTPSSGJTLNKLCTGTASDLOEATYFEPDSPEGIDNNSSLEHTPM
 AGDTHSGGMPVTPSSGJTLNKLCTGTASDLOEATYFEPDSPEGIDNNSSLEHTPM
 NIOBSMGKANSWTEGSLFKGKAHGPALITKDNALRYVSLITLSENNDSKLESGIM
 KTHIDGPELTPNSPYWONITLSDPTERKATVPLIHDMLIDNKATLRTLNHSNKT
 NTHOEXKITLEIKETKEPTLQENWALPOLHETVQKGNKMKLITLSTROUVENSDYGA
 YAPULOCIRSEGOEPTLEKRYVKGSEFTDVOJEMVPPSSRNPLTNLNDHN
 SOONVFORSKALKOPRLEJLETELEKRIYDPTSSNKNMHLKOTVEKATYATNITSPNT
 KEKALATGOSLCTLRHSITLQOARNSPLKAYASPSIRIYATLVTFONSSHLR
 AASFKKDSOSLSEHSHFLOGAKRNILSLATITLMTDQKVESELOSTASNTSVYK
 ENYLVKPDLPRTSGVVELLPVHLYOKDLPPTETLSNGSHGDLVEGSLTOSPEKA
 KKNEDBNRPKVPFLRAVETASKIPSKLIDPLAMDNYHGTQJLPERKRSOESKREKA
 FKKDITLILANCSNNAITAIINSGCONPELEVTAMWAKGTETLCONPULTRHORE
 ITRTTLQSOOEIDYDPTIYEMKREKPELEVTAMWAKGTETLCONPULTRHORE
 EVDENMPTVFNRAOSRPSYSLSTLSEEDORGAAPRKNYKPNETATYKMWOLLYIA
 IADKREPCOKMAFYSDVLDLEKYNHSLIPLELVCYTNLNDNARIHROTVYDOPALFT
 IADKREPCOKMAFYSDVLDLEKYNHSLIPLELVCYTNLNDNARIHROTVYDOPALFT
 IAWTILSNGSNHTLHFSHGQCPYRKAEKMLALINJGVETVEMJPSAGIWR
 VECILGHEHJAGNHTLHFSHGQCPYRKAEKMLALINJGVETVEMJPSAGIWR
 SOSTINASTEPFMSIVDLAPMIIHGIKTGQASORQSLISQIOTIAMPKCLARIY
 TYRONSJSTLMEFFGVNDSGIKRINENPIITARIYLRPHYTSRISTLMELMGDZL

Query Match	57.5%;	Score 575;	DB 6;	Length 8967.
Best Local Similarity	100.0%;	Pred. No. 5.2e-164;		
Matches 575; Conservative	0. w.i.			

486 3211- 203 CAGCAGAAATAGAGCTCTCCACCTGCTTCTTCTGTGCTTTTGGGATTCGCTTTATCCG 160

169 CACCAGAACATCGTCTGCGCTGCACTGTCATGGACTATATGCAAGTGATCT 545

546 cggtagactgcctatgaacgaaattttt
228

CGGTGACTGCCTGTGACCGCAATATTTCTTCCATTCaa 605

606 caccctcagtcgtatcaaaaagactctgttttgaagaattcaacatt
|||||

289 CACCTCAGTCGTGATCAAAAAAGACTCTGTTGTAGATTCCGGATACCmmmmmmcat 665

349 GCGGCTGGTGTCTGCTAGTCTCATCATTCCAGGCTGAATTTA TTT

726 tqatacatnctacttt
 ccccccclggatgsgtctgctaagtcctaccattccaggctgaggttta 408

[illegible]

786 tgggtatctactggaagcttctgaaggaatctcttcttccacacccatcttcacgtcttcacgtctgt 468

469 TGGTGTATCCCTACTGGAACCTTTCTGAGGGAGCTGAATTATGCATC...
845

040 ygaagaagaatgatcaagtcctccctgtygaagccatacatatgctcctgaacat- 328
|||||

906 aaagctcttccttgtagaccatcatatgctgcaggtctt 588

|||||
|||cccaatggcctctgaccacactgtgccttacctactcatactttctca 965
589 GAAGAGGATGTCTCCATTCCCA

966 tctgacctgtataaaagacttgaattat
 caccacaccccttacctaactatattcttcaca 648

649 TGTGGACCTGTAAAGACTTGAATTCAGGCTCA 603

12

171409 8967 hz

03-APR-1998

Unknown
01:300/344

Unknown.
Unclassified.

1 (bases 1 to 8967)
HORS
Bodner, M., De Polo, N.J., Chang, C.

Patent: US 5681746-A 1 28-OCT-1997.

1. .8967
/Organism-Plant

Year	1898	1903	1905
a	2841	2841	2841
c	1898	1898	1898
g	1833	1833	1833
t	2395	2395	2395

[illegible]

[illegible]

5

1

XX A DNA molecule (AA169811) codes for Factor VIII-dB695-HCII (AA186555) a hybrid protein in which amino acids 712-736 of Factor-dB695 (Factor VIII (de1868-1562) B-domain are replaced by amino acids 51-80 from the acidic region) and potential thrombin-binding site of human heparin cofactor II (HCII). It was obtained by PCR amplification (see also AA169812-13) of the HCII acidic region from CC liver cDNA, fusion to sequences encoding Factor VIII CC aa706-711 and aa737-743, and incorporation of the construct into CC plasmid pCDB-dB695. The hybrid protein, which can be expressed CC using gene therapy techniques, has increased procoagulant activity CC owing to the HCII acidic region, and can be used to treat blood CC coagulation disorders such as haemophilia A.

XX Sequence 5035 bp: 1484 bp: 1300 bp: 1100 bp: 900 bp: 700 bp: 500 bp: 300 bp: 100 bp: 0 bp

Query Match	57.7%;	Score 577.4;	DB 18;	Length 5035;
Best Local Similarity	99.8%;	Pred. No. 3.9e-156;		
Matches 578; Conservative	0;	Mismatches 1		

Db 30 ccaccacgagaaatagagctctccacctgctctcttcctctgctcttcttgagattcgattcgattta 4
Qy 482 gtgcacaccaaataatactaccggtgtgacgtgagaaactgatactaggagactatagcaaaagt 5
Db 90 gtgcacaccaaagaatacaactactggtgtgacgtgagaaactgatactaggagactatagcaaaagt 8
Qy 542 atctcggtgagctgtgcctgtgtgacgcaagattctctctctgagtgccaaatcttttccat 1
Db 150 atctcggtgagctgtgcctgtgtgacgcaagattctctctctgagtgccaaatcttttccat 6
Qy 602 tcaaacacttcagctgtgtataaaaaaacctctgtttgtagaaatacagatcacacttttca 2
Db 210 tcaaacacttcagctgtgtataaaaaaacctctgtttgtagaaatacagatcacacttttca 6
Qy 662 acatgcgttaagccaagagccacacctgagatcggtgcgtcaggatctccacatccagctgaggg 26
Db 270 acatgcgttaagccaagagccacacctgagatcggtgcgtcaggatctccacatccagctgaggg 72
Qy 722 ttattatatacagtggtcatcttaacttaagaaacagatcgctccacatccctgtcagatcttcaatg 32
Db 330 ttattatatacagtggtcatcttaacttaagaaacagatcgctccacatccctgtcagatcttcaatg 388

OY 782 ctgtgtgtatcctactggaagctcttgagagctgataatgatacagaccagtc 841
 |||||||
 Db 390 ctgtgtgtatcctactggaagctcttgagagctgataatgatacagaccagtc 449
 OY 842 aaagggagaagaagatgataagctcttcctcgtgtgaaagccatactatgtctgcag 901
 |||||||
 Db 450 aaagggagaagaagatgataagctcttcctcgtgtgaaagccatactatgtctgcag 509
 OY 902 tccggaagaagaatggtccaaatggtcctcgtgaccacactgtgcttactactatctt 961
 |||||||
 Db 510 tccggaagaagaatggtccaaatggtcctcgtgaccacactgtgcttactactatctt 569
 OY 962 ctcaatgtgacctgtgataaagacttgatcagagctca 1000
 |||||||
 Db 570 ctcaatgtgacctgtgataaagacttgatcagagctca 608
 RESULT 7
 ID AAA49232 standard; DNA; 12445 BP.
 AC AAA49232:
 XX 26-SEP-2000 (first entry)
 DT Vector HSQReNeo for transforming endothelial cells.
 DE
 XX Hemostatic; antiplatelet; antidiabetic; neuroprotective; vector;
 KW osteoporosis; antistickling; immunostimulant; gene therapy; collagen;
 KW endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;
 KW vascular endothelial growth factor; bovine brain extract; haemophilia;
 KW Factor VIII; human; thalassemia; adenine deaminase deficiency; ss;
 KW sickle cell anemia; thalassemia; diabetes; alpha antitrypsin deficiency;
 KW Alzheimer's disease; brain disease; heart disease; immune system defect;
 KW bone fracture; osteoporosis.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200032750-A1.
 XX 08-JUN-2000.
 PD
 XX 24-NOV-1999; 99MO-US28033.
 PF
 XX 24-NOV-1998; 98US-0109687.
 PR
 XX (MINU) UNIV MINNESOTA.
 PA (UTEM) UNIV EMORY.
 PA (HEBB) HEBBEL R P.
 PA (LINY) LIN Y.
 PA (LOLL) LOLLAR J S.
 XX
 XX Hebbel RP, Lin Y, Lollar JS;
 PI
 DR MPI: 2000-412303/35.
 XX
 XX Expanding population of endothelial cells useful to biocompatibilize
 PT implantable medical devices comprises contacting buffy coat cells with
 PT collagen I coated surface in culture medium comprising vascular
 PT endothelial growth factor
 XX
 PS Claim 19; Fig 4; 53pp; English.
 XX
 XX The invention relates to a method for expanding the population of
 CC endothelial cells (EC) obtained from peripheral blood, by culturing,
 CC in contact with a collagen I coated surface, buffy coat cells obtained
 CC from peripheral mammalian blood in the presence of a culture medium
 CC containing vascular endothelial growth factor (VEGF) and free of bovine
 CC brain extract. EC are useful for treating hemophilia by introducing it
 CC into the blood stream of a mammal, so that an effective amount of
 CC Factor VIII protein is secreted in the blood stream of the mammal.

CC This sequence represents the vector HSQReNeo where the human factor
 CC VIII gene into which an enhanced green fluorescent protein coding
 CC sequence (HSQ) has been inserted, is subcloned. Transgenic EC transduced
 CC in vitro are useful for improving prosthetic implants. EC is also useful
 CC for diagnosing clotting disorders where indication of an enzyme. EC is also
 CC associated with a reduction in the activity of an enzyme. EC is also
 CC useful in gene therapy for treating the variety of diseases including
 CC adenine deaminase deficiency, sickle cell anemia, thalassemia,
 CC hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders
 CC such as Alzheimer's disease, heart diseases, defects in immune system,
 CC for repairing bone fractures and to treat or prevent osteoporosis.
 CC
 XX Sequence 12445 BP; 3255 A; 2989 C; 2951 G; 3250 T; 0 other;
 SO

Query Match 57.7%; Score 577.4; DB 21; Length 12445;
 Best Local Similarity 99.8%; Pred. No. 5.8e-156;
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 422 ccacacagcaaatagagctctcactgctctctctctgtgctcttgccttgcgattcgtctta 481
 |||||||
 Db 763 ccacacagcaaatagagctctcactgctctctctctgtgctcttgcgattcgtctta 822
 OY 482 gtgcacacagcaaatagagctctcactgctctctctctgtgctcttgcgattcgtctta 882
 |||||||
 Db 823 gtgcacacagcaaatagagctctcactgctctctctctgtgctcttgcgattcgtctta 601
 OY 542 atctcgtgtgagctgtcgtgagcagcaagattcctcctcctcctcctcctcctcctcctcct 942
 |||||||
 Db 883 atctcgtgtgagctgtcgtgagcagcaagattcctcctcctcctcctcctcctcctcctcct 661
 OY 602 tcaacacacagcaaatagagctctcactgctctctctctgtgctcttgcgattcgtctta 1002
 |||||||
 Db 943 tcaacacacagcaaatagagctctcactgctctctctctgtgctcttgcgattcgtctta 721
 OY 662 acatcgtgtgagctgtcgtgagcagcaagattcctcctcctcctcctcctcctcctcctcct 1062
 |||||||
 Db 1003 acatcgtgtgagctgtcgtgagcagcaagattcctcctcctcctcctcctcctcctcctcct 781
 OY 722 ttatgatacagtggtatcacttaagaacatggtcctcctcctcctcctcctcctcctcctcct 1122
 |||||||
 Db 1063 ttatgatacagtggtatcacttaagaacatggtcctcctcctcctcctcctcctcctcctcct 841
 OY 782 ctgtgtgtatcctactggaagctcttgagagctgataatgatacagaccagtc 1182
 |||||||
 Db 1123 ctgtgtgtatcctactggaagctcttgagagctgataatgatacagaccagtc 901
 OY 842 aaagggagaagaagatgataagctcttcctcgtgtgaaagccatactatgtctgcag 1242
 |||||||
 Db 1183 aaagggagaagaagatgataagctcttcctcgtgtgaaagccatactatgtctgcag 961
 OY 902 tccggaagaagaatggtccaaatggtcctcgtgaccacactgtgcttactactatctt 1302
 |||||||
 Db 1243 tccggaagaagaatggtccaaatggtcctcgtgaccacactgtgcttactactatctt 1000
 OY 962 ctcaatgtgacctgtgataaagacttgatcagagctca 1341
 |||||||
 Db 1303 ctcaatgtgacctgtgataaagacttgatcagagctca

RESULT 8
 ID AAX82259 standard; cDNA; 9164 BP.
 AC AAX82259:
 XX
 XX 18-AUG-1999 (first entry)
 DT
 XX
 XX Beta-domain deleted Factor VIII protein encoding gene (uncorrected).
 DE
 XX Factor VIII protein; gene modification; gene therapy; clinical disorder;
 KW Factor VIII protein; gene modification; gene regulation; beta-domain; human;
 KW splicing pattern; RNA processing; gene regulation; beta-domain; human;
 KW ss.

```

XX OS Homo sapiens.
XX PN MO9929848-A1.
XX PD 17-JUN-1999.
XX PF 25-NOV-1998; 98WO-US25354.
XX PR 16-JAN-1998; 98US-0071596.
XX PR 05-DEC-1997; 97US-0067614.
XX PA (IMMU-) IMMUNE RESPONSE CORP.
XX PI Bidlingmaier S, Gonzales JEN, ILL CR, Yang CQ;
XX DR WPI: 1999-385602/32.
XX DR P-PSDB; AAY21675.
XX PT Genes and vectors exhibiting increased expression and novel splicing
XX PT patterns, useful for expression of, e.g. beta-domain deleted factor
XX PS VIII
XX PS Example 2; Page 79-89; 123pp; English.
CC The invention describes novel genes and vectors exhibiting increased
CC expression and novel splicing patterns. It provides a gene encoding a
CC Factor VIII protein, that comprises one or more consensus or near
CC consensus splice sites which have been corrected to increase expression.
CC The method, DNA sequences and expression vectors can be used to increase
CC the expression of a gene, especially a Factor VIII gene. Genes containing
CC modified 5' and/or 3' untranslated regions have optimized expression
CC levels and tissue-specific expression. The methods are used for
CC identification and correction of consensus splice sites, addition of
CC introns, optimization of 5' and 3' untranslated regions and increase in
CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy
CC to treat a clinical disorder. to study RNA processing and/or gene
CC regulation. The present sequence represents an uncorrected version of
CC beta-domain deleted factor VIII protein encoding gene sequence (construct
CC PCY-2). This was used to develop a new coding sequence for beta-domain
CC deleted factor VIII protein by correcting the consensus splice sites.
XX SO Sequence 9164 BP; 2451 A; 2170 C; 2081 G; 2462 T; 0 other;

Query Match          57.6%; Score 575.8; DB 20; Length 9164;
Best Local Similarity 99.7%; Pred. No. 1.3e-155;
Matches 577; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 422 ccaccatgaataaagcttcacactgtctcttctgtgcttggagtttgcctta 481
DB 1001 ccaccatgaataaagcttcacactgtctcttctgtgcttggagtttgcctta 481
OY 482 gtgcacacagaagaatacactgtgagtgagtgagtgagtgagtgagtgagtgag 1060
DB 1061 gtgcacacagaagaatacactgtgagtgagtgagtgagtgagtgagtgagtgag 541
OY 542 atctcggtgagctcctgtgagcgaagatttccctcctgagtgagtgagtgagtgag 1120
DB 1121 atctcggtgagctcctgtgagcgaagatttccctcctgagtgagtgagtgagtgag 601
OY 602 tcaacacctagtcgtgtacaaaagaactcttctgtgagtgagtgagtgagtgagtgag 1180
DB 1181 tcaacacctagtcgtgtacaaaagaactcttctgtgagtgagtgagtgagtgagtgag 661
OY 662 acatcgctgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaag 1240
DB 1241 acatcgctgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaag 721
OY 722 ttatgatagcagtggtcatlaaacttaagaacatgaggtccatcctgtgaagcttcacatg 1300
DB 1301 ttatgatagcagtggtcatlaaacttaagaacatgaggtccatcctgtgaagcttcacatg 1360

```

```

OY 782 ctgttggtgatactactggaagaagctctgaggagcgtgaatatgatcatcagaccagtc 841
DB 1361 ctgttggtgatactactggaagaagctctgaggagcgtgaatatgatcatcagaccagtc 1420
OY 842 aaaggaggaagaagaatgaataagcttccctggtgaggaagcacaatataatgtctgcagc 901
DB 1421 aaaggaggaagaagaatgaataagcttccctggtgaggaagcacaatataatgtctgcagc 901
OY 902 tccctgaagaagaatggtcccaatgagcctctgacccacgtgtgcttaccatcatcttc 961
DB 1481 tccctgaagaagaatggtcccaatgagcctctgacccacgtgtgcttaccatcatcttc 961
OY 962 ctcatgtgacctgtgtaaaagactgaattcaggcccca 1000
DB 1541 ctcatgtgacctgtgtaaaagactgaattcaggcccca 1579

RESULT 9
AAK82261
ID AAK82261 standard; cDNA; 11846 BP.
XX AC AAK82261;
XX DT 18-AUG-1999 (first entry)
XX DE Factor VIII protein encoding gene (construct pLZ-6).
XX KW Factor VIII protein; gene modification; gene therapy; clinical disorder;
XX KW splicing pattern; RNA processing; gene regulation; beta-domain; human;
XX SS.
XX OS Homo sapiens.
XX PN MO9929848-A1.
XX PD 17-JUN-1999.
XX PF 25-NOV-1998; 98WO-US25354.
XX PR 16-JAN-1998; 98US-0071596.
XX PR 05-DEC-1997; 97US-0067614.
XX PA (IMMU-) IMMUNE RESPONSE CORP.
XX PI Bidlingmaier S, Gonzales JEN, ILL CR, Yang CQ;
XX DR WPI: 1999-385602/32.
XX DR P-PSDB; AAY21676.
XX PT Genes and vectors exhibiting increased expression and novel splicing
XX PT patterns, useful for expression of, e.g. beta-domain deleted factor
XX PT VIII
XX PS Example 1; Page 101-115; 123pp; English.
CC The invention describes novel genes and vectors exhibiting increased
CC expression and novel splicing patterns. It provides a gene encoding a
CC Factor VIII protein, that comprises one or more consensus or near
CC consensus splice sites which have been corrected to increase expression.
CC The method, DNA sequences and expression vectors can be used to increase
CC the expression of a gene, especially a Factor VIII gene. Genes containing
CC modified 5' and/or 3' untranslated regions have optimized expression
CC levels and tissue-specific expression. The methods are used for
CC identification and correction of consensus splice sites, addition of
CC introns, optimization of 5' and 3' untranslated regions and increase in
CC cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy
CC to treat a clinical disorder. to study RNA processing and/or gene
CC regulation. The present sequence represents the full length factor VIII
CC protein encoding gene sequence (construct pLZ-6). This contains an
CC intron spanning the beta-domain.
XX SO Sequence 11846 BP; 3419 A; 2718 C; 2607 G; 3102 T; 0 other;

```


QY 782 ctgtgtgtatcctactggaagctcttgaggagctgaatataatgatacgaaccagtc 841
 Db 1361 ctgtgtgtatcctactggaagctcttgaggagctgaatataatgatacgaaccagtc 1420
 QY 842 aaaggaggaagagatgataaagctcttcctgtgtggaagcatatcttcggcagg 901
 Db 1421 aaaggaggaagagatgataaagctcttcctgtgtggaagcatatcttcggcagg 1480
 QY 902 tcttgaaagagatgttccatgtgctctgaccacgtgcttaccatcatatctt 961
 Db 1481 tcttgaaagagatgttccatgtgctctgaccacgtgcttaccatcatatctt 1540
 QY 962 ctcatgtgacctgtgaaagacttgaattcaagcctca 1000
 Db 1541 ctcatgtgacctgtgaaagacttgaattcaagcctca 1579

RESULT 11

AAV19581
 ID AAV19581 standard; cDNA; 4832 BP.

AAV19581;

06-AUG-1998 (first entry)

Human factor VIII beta-domain deleted SQN deletion cDNA sequence.

Replication defective; recombinant retrovirus; RV; therapeutic protein; haemophilia; thrombosis; hypercoagulable disorder; liver disease; human; cystic fibrosis; Duchenne's muscular dystrophy; Lesch-Nyhan syndrome; diabetes; hypopituitarism; adenine deaminase deficiency; hypercholesterolemia; Guader's syndrome; high blood pressure; Alzheimer's disease; anaemia; inflammatory disease; factor VIII; ss.

Homo sapiens.

MO9800541-A2.

08-JAN-1998.

02-JUL-1997; 97WO-US11784.

04-JUN-1997; 97US-0869309.

03-JUL-1996; 96US-0645601.

13-AUG-1996; 96US-0696381.

(CHIR) CHIRON CORP.

Allen JR, Barber JR, Boder M, Chang SMW, Chong K;
 De LA VEGA D, Depoloni, Greenard J, Hsu DC, Ibanez CE;
 Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;
 WPI: 1998-086966/08.

P-PSDB: AAM6246.

New replication defective recombinant retro-viruses - which can be administered to provide long term systemic expression of therapeutic protein in blood, useful in, e.g. treating hyper-coagulable disorders

Example 28; Pages 210-213; 272pp; English.

This cDNA encodes the beta-domain deleted SQN deletion protein of human factor VIII. This is used in the construction of recombinant retroviral vectors expressing human factor VIII. The invention provides the preparation of replication defective recombinant retrovirus (RV) expressing a therapeutic protein. The RV preparation is resistant to degradation by human complement and is capable of inducing long term systemic expression of the therapeutic protein when administered intravenously to a human. The long term systemic expression results in a measurable level of the therapeutic protein being produced in the blood of the human for a period of at least 30 days after the administration of

the RV vector preparation. RV's can be used for in vivo delivery of therapeutic protein to treat, e.g. haemophilia A, haemophilia B, thrombosis, hypercoagulable disorders, liver diseases such as hepatitis, severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's muscular dystrophy, inherited emphysema, familial hypercholesterolemia, antitrypsin deficiency, Guader's syndrome, anaemia, infections such as HIV infection, high blood pressure, Alzheimer's disease, autoimmune or inflammatory disease or graft versus host disease. RV's are capable of surviving inactivation in human serum thereby allowing efficient gene transfer over prolonged periods of time.

Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 57.5%; Score 575; DB 19; Length 4832;
 Best Local Similarity 100.0%; Pred. NO. 1.9e-155;

Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 426 catgcaatagagctctccactgctcttctgtgtgcttgcgattcgcttagtgc 485
 Db 71 catgcaatagagctctccactgctcttctgtgtgcttgcgattcgcttagtgc 130
 QY 486 caacgaagatctactcctggtgtgagtgagtgatgagatgagatgagatgagat 545
 Db 131 caacgaagatctactcctggtgtgagtgagtgatgagatgagatgagatgagat 190
 QY 546 cgttgagtgcttcttgagcgaagattctctctctagagtgccaaatcttccattca 605
 Db 191 cgttgagtgcttcttgagcgaagattctctctctagagtgccaaatcttccattca 250
 QY 606 caactcagtcgtgtacaaaagactcgtgtgtgagatgagatgagatgagatgagat 665
 Db 251 caactcagtcgtgtacaaaagactcgtgtgtgagatgagatgagatgagatgagat 310
 QY 666 cgtcaagcaagacccctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 725
 Db 311 cgtcaagcaagacccctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 370
 QY 726 tgatacagtggtatctacattcaagaacatgagtgctccatctgtcagttcagtcgt 785
 Db 371 tgatacagtggtatctacattcaagaacatgagtgctccatctgtcagttcagtcgt 840
 QY 786 tgggtatcctactggaagctcttgaggagctgaatagatgatacgaagcagtcag 845
 Db 431 tgggtatcctactggaagctcttgaggagctgaatagatgatacgaagcagtcag 905
 QY 846 gtagaagaagaatgataagctctccctggtgtgagagcatatgcttgcagtgct 905
 Db 491 gtagaagaagaatgataagctctccctggtgtgagagcatatgcttgcagtgct 550
 QY 906 gaaagaagatgtccatgagctctgacacgtgtgcttaccatcatatcttccca 965
 Db 551 gaaagaagatgtccatgagctctgacacgtgtgcttaccatcatatcttccca 610
 QY 966 tgtgacctgtgtaaaagacttgaattcaagcctca 1000
 Db 611 tgtgacctgtgtaaaagacttgaattcaagcctca 645

RESULT 12

AAV15338
 ID AAV15338 standard; DNA; 4832 BP.

AAV15338;

20-JUL-1998 (first entry)

Human Factor VIII SQN deletion mutant DNA.

Factor VIII; blood clotting; haemophilia A; gene therapy;
 retrovirus; vector; human; ss.

XX	Homo sapiens.
OS	Synthetic.
XX	Location/Qualifiers
FH	Key
FT	CDS
PN	72..4445 /*tag= a
PD	W09800542-A2.
PP	08-JAN-1998.
PX	02-JUL-1997; 97WO-US11785.
PR	04-JUN-1997; 97US-0869309. 03-JUL-1996; 96US-0645601. 13-AUG-1996; 96US-0696381.
PA	(CHIR) CHIRON CORP.
PI	Allen JR., Barber JR., Bodder M., Chang SM., Chong K; De La Vega D., Depolo NJ., Greengard J., Hsu DC., Ibanez CE; Jolly DJ., Mittelstaedt DM., Prussak CE., Respess JG;
PT	New replication defective recombinant retroviruses - which express B domain-deleted human factor VIII or human factor IX for the treatment of haemophilia
DR	WPI: 1998-086967/08. P-PSTDB: AAM44372.
PS	Claim 6; Page 174-175; 236pp; English.
CC	This DNA sequence includes a coding region for the B domain deletion mutant SQN (see AAM44372) of human Factor VIII. The SQN mutant is created by fusing Ser-743 to Gln-1638 of native Factor VIII (see AAM44373) to form a Ser-Gln-Asn (SQN) link between the A2 and A3 Factor VIII domains. When compared to plasmid-derived Factor VIII, the SQN deletion does not influence the in vivo pharmacokinetics, but the reduced size of the molecule appears to decrease proteolytic degradation. The invention relates to preparations of replication defective recombinant retrovirus (RV) expressing a B domain deleted human Factor VIII protein, where the recombinant RV is capable of infecting human cells, is resistant to degradation by human complement and is capable of inducing long-term (at least 30 days and up to 6 months or longer post-infection) systemic expression of Factor VIII when administered to a haemophilia A patient.
SC	Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other:
Query Match	57.5%; Score 575; DB 19; Length 4832;
Best Local Similarity	100.0%; Pred. NO. 1.9e+155;
Matches	575; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY	426 catgcaaatagagtctccacacctgtccttcttgatggccttttcgatcttgtcttatgc 485
DB	71 catgaataatagaactctccaacctgtcttcttgatggccttttcgatcttgtcttatgc 130
OY	486 caccagaagatatctaactcggtgtagcgaatgcatcatgaagatatgc 545
DB	131 caccagaagatatctaactcggtgtagcgaatgcatcatgaagatatgc 190
OY	546 cggtgagctgccctgtgtagcgaagatttccctoctagatgagcacaatctttccatcaa 605
DB	191 cggtagcgtccctgtgtagcgaagatttccctoctagatgagcacaatctttccatcaa 250
OY	606 cacctcacgtcgtgtacaaaagactctgtttgtagaaatltaagratoccttticaaac 665
DB	251 cacctcacgtcgtgtacaaaagactctgtttgtagaaatltaagratoccttticaaac 310
OY	666 cgctgaagccaaggccacctgtgtagtctctagtgltccttacatlcacgtctgagttaa 725

Db	311	cgctaaagcccaagccaccctgcatgggtctgctaaagttccacacccaaggtctgaagttta	370
Oy	726	tgatacagtggtcattacacttaagaacatggtcttccacatcctgtcagtcttcaatgctgt	785
Db	371	tgatacagtggttcatctacacttaagaacatggtcttccacatcctgtcagtcttcaatgctgt	430
Oy	786	tgtgtatctcacttgtaaaagctctcttgaggagctcggaatagatgatacagaccagttcaag	845
Db	431	tgtgtatctcacttgtaaaagctctcttgaggagctcggaatagatgatacagaccagttcaag	490
Oy	846	ggagaagaagaatgataatgaagttctccctctggttggaagccatacatatgcttgagagttct	905
Db	491	ggagaagaagaatgataatgaagttctccctctggttggaagccatacatatgcttgagagttct	550
Oy	906	gaaagaagaatggtccaaatgagctctcgaccacactgtgcttaccatccatcatattcttca	965
Db	551	gaaagaagaatggtccaaatgagctctcgaccacactgtgcttaccatccatcatattcttca	610
Oy	966	tgttgacctgtgtaaaagacttgaaattcaagcctca	1000
Db	611	tgttgacctgtgtaaaagacttgaaattcaagcctca	645
RESULT	13		
AAr03571		standard; cDNA: 6300 BP.	
ID	AAr03571		
XX	AAr03571;		
AC			
XX			
XX	13-MAY-1996	(first entry)	
DT			
XX			
DE	Factor-VIII cDNA.		
XX			
KW	Factor-VIII; blood clotting; haemophilia A; ss.		
XX			
OS	Homo sapiens.		
XX			
EH	Location/Qualifiers		
XX	Key	7..6300	
EH	CDS	/*tag= a	
FT		7..63	
FT	sig_peptide	/*tag= b	
FT		64..6300	
FT	mat_peptide	/*tag= C	
FT		/note= "the sequence is incomplete since	
FT		it encodes only amino acids 1-2079	
FT		of mature Factor-VIII"	
FT			
FT			
XX	EP690126-A1.		
XX			
PD	03-JAN-1996.		
XX			
PF	13-JUN-1988;	88EP-0110466.	
XX			
PR	12-JUN-1987;	87EP-0201121.	
XX			
PA	(IMMO) IMMO AG.		
XX			
XX			
PI	Pannekoek H, Van Leen RW, Van Ooyen AJJ, Verbeet MP;		
XX			
DR	WPI: 1996-051261/06.		
XX			
DR	P-PSDB: AAR6863.		
XX			
PT	New deletion mutant of factor VIII - useful for treating factor VIII		
XX	deficiency, esp. haemophilia A		
XX			
XX	Example 1; Fig 1; 31pp; English.		
XX			
CC	A human liver cDNA sequence (AAr03571) in expression vector pCLB89		
CC	encodes the Factor-VIII precursor (see AAR6863). The cDNA was		
CC	subcloned to loop-out mutagenesis and used in the construction of		
CC	vectors utilised in the transformation of e.g. COS, CHO and C127		

QY 906 gaaagaatgtgtccatgtgctctgaccacgtgtgcttaactaataatcttctca 965
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 486 gaaagaatgtgtccatgtgctctgaccacgtgtgcttaactaataatcttctca 545
 QY 966 tgtgacctgtgttaagaacttgaattcaagcctca 1000
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 546 tgtgacctgtgttaagaacttgaattcaagcctca 580

RESULT 15

AA81096
 ID AA81096 standard; cDNA; 8241 BP.

AC AA81096;

DT 29-OCT-1990 (first entry)

DE cDNA sequence encoding human Factor VIII signal sequence and mature
 DE protein on plasmid pCLB89.

DE Haemophilia A; bloodclotting; treatment; diagnosis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 7..63 /*tag= a

XX mat_peptide /product=Signal peptide

XX 64..7062

XX /*tag= b

XX EP294910-A.

XX 14-DEC-1988.

XX 13-JUN-1988; 88EP-0201209.

XX 12-JUN-1987; 87EP-0201121.

XX 03-JUN-1988; 88EP-0201209.

XX (KONN) GIST-BROCADES NV.

XX van Ooyen AJJ, Pannekoek H, Verbeet MP, van Leen RW;

XX WPI: 1988-355361/50.

XX P-PSDB; AAP80659.

XX proteins having Factor VIII activity -

XX comprising deletion mutant proteins of Factor VIII in which

XX central region has been deleted

XX Example; Fig 1; 37pp; English.

XX The isolation of Factor VIII mRNA from human liver, and the preparation,
 CC purification and identification of cDNA and its assembly in the plasmid
 CC pEP121 resulting in plasmid pCLB89 have been described in patent
 CC application EP 0253455. The Factor VIII polypeptides of the invention
 CC include deletion mutant proteins of Factor VIII in which the central
 CC region or "B domain" as well as a portion of the 92 kD region has been
 CC deleted. The polypeptides have enhanced Factor VIII activity and/or
 CC decreased immunogenicity and can be used for the treatment of
 CC Haemophilia A. They can also be used to prepare antibodies. The Abs.
 CC can be used to determine the concn. of Factor VIII in a body fluid.

CC Sequence 8241 BP; 2625 A; 1759 G; 1703 C; 2154 T; 0 other;

QY 426 catgcaaatagagctctccacctgtctcttctgtgcttcttgcatctgcttaagtc 485

Query Match 57.5%; Score 575; DB 9; Length 8241;
 Best Local Similarity 100.0%; Pred. No. 2.4e-155; Indels 0; Gaps 0;
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6 catgcaaatagagctctccacctgtctcttctgtgcttcttgcatctgcttaagtc 65
 QY 486 caccagaagaatactacctgtgtgtgcaatgtgcatggtgactatagcaaatgtact 545
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 66 caccagaagaatactacctgtgtgtgcaatgtgcatggtgactatagcaaatgtact 125
 QY 546 cgtgtgacctgtgtgtgcaatgtgcatggtgactatagcaaatgtact 605
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 126 cgtgtgacctgtgtgtgcaatgtgcatggtgactatagcaaatgtact 185
 QY 606 cactcagctgtgtgtgcaatgtgcatggtgactatagcaaatgtact 665
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 186 cactcagctgtgtgtgcaatgtgcatggtgactatagcaaatgtact 245
 QY 666 cgtgtgacctgtgtgtgcaatgtgcatggtgactatagcaaatgtact 725
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 246 cgtgtgacctgtgtgtgcaatgtgcatggtgactatagcaaatgtact 305
 QY 726 tgatacagtggtcattacttaacttaagaacatgtggtccatcctgtcaagcttcagtcgt 785
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 306 tgatacagtggtcattacttaacttaagaacatgtggtccatcctgtcaagcttcagtcgt 365
 QY 786 tgggtatcttacttggaagcttctggaagagctggaatgatgatcagaaccagtcgaag 845
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 366 tgggtatcttacttggaagcttctggaagagctggaatgatgatcagaaccagtcgaag 425
 QY 846 ggaagaagaagaatgataaagcttctcctgtgtggaagcacaataatgtcttggaggtctct 905
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 426 ggaagaagaagaatgataaagcttctcctgtgtggaagcacaataatgtcttggaggtctct 485
 QY 906 gaaagaatgtgtccaaatgtgctctgaccacgtgtgcttaactaataatcttctca 965
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 486 gaaagaatgtgtccaaatgtgctctgaccacgtgtgcttaactaataatcttctca 545
 QY 966 tgtgacctgtgttaagaacttgaattcaagcctca 1000
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 546 tgtgacctgtgttaagaacttgaattcaagcctca 580

Search completed: January 17, 2002, 16:32:59
 Job time: 12484 sec

Tue Jan 22 15:12:26 2002

us-09-740-211-13_copy_1_1000.png

[illegible]

Db	241	aaagaagtatattagaagcagcagctcttccttcgcacacagaataccttccggcgtgcgcgcct	3000
OY	301	aggcaggtiaagtcgctgtgtgtgtcccgaggccgcgcctctcttaagggtatagccct	360
Db	301	aggcaggtiaagtcgctgtgtgtgtcccgaggccgcgcctctcttaagggtatagccct	360
OY	361	tcgtgtgccttgaattactgtaacactgtaacatccaaattttcttttcccaagttatgat	420
Db	361	tcgtgtgccttgaattactgtaacactgtaacatccaaattttcttttcccaagttatgat	420
OY	421	tcaccacagcaataaagaagctctccaccgctcttctctctgtgccttcttgatctgcctt	480
Db	421	tcaccacagcaataaagaagctctccaccgctcttctctctgtgccttcttgatctgcctt	480
OY	481	agtcgcacacagaagatactacccctgggtggaagtcgaacgtgcataatggagataatgcaagt	540
Db	481	agtcgcacacagaagatactacccctgggtggaagtcgaacgtgcataatggagataatgcaagt	540
OY	541	gattcgtgtagctgcctgtgtgagcgaagatttccctccctagaagtcgccaatatcttcca	600
Db	541	gattcgtgtagctgcctgtgtgagcgaagatttccctccctagaagtcgccaatatcttcca	600
OY	601	ttcaaaacctcagtcgtgtgtacaaaagaactctgtcttgaatctcaaggtatcactcttc	660
Db	601	ttcaaaacctcagtcgtgtgtacaaaagaactctgtcttgaatctcaaggtatcactcttc	660
OY	661	aaacatgcgtcaaacgaagcgcacaccttggaatgggtcctcctagctccatacatccagctgag	720
Db	661	aaacatgcgtcaaacgaagcgcacaccttggaatgggtcctcctagctccatacatccagctgag	720
OY	721	gtttatgatacagtggtgtcatttaacactaaagaacatggtctccacacgtctgaagcttcat	780
Db	721	gtttatgatacagtggtgtcatttaacactaaagaacatggtctccacacgtctgaagcttcat	780
OY	781	gctgtgtgtgtatccctacttggaaagcttcttgggggaagcttgataatgatatgacagcaagt	840
Db	781	gctgtgtgtgtatccctacttggaaagcttcttgggggaagcttgataatgatatgacagcaagt	840
OY	841	caaaaggagaaagaagaatgtataaagcttccctctgtgtgaaagccataatattgtctggag	900
Db	841	caaaaggagaaagaagaatgtataaagcttccctctgtgtgaaagccataatattgtctggag	900
OY	901	gtctctgaagaagaatgttcccaatggcctcttgcacacactgtgccttatacctaactatctt	960
Db	901	gtctctgaagaagaatgttcccaatggcctcttgcacacactgtgccttatacctaactatctt	960
OY	961	tctcatgtgacctgtttaaagaacttgaattcaagctca	1000
Db	961	tctcatgtgacctgtttaaagaacttgaattcaagctca	1000

```

RESULT 2
US-09-364-862-13
; Sequence 13, Application US/09364862
; Patent No. 6221349
GENERAL INFORMATION:
; APPLICANT: Coto, Linda B.
; APPLICANT: Colosi, Peter C.
; TITLE OF INVENTION: ADEMO ASSOCIATED VECT
; TITLE OF INVENTION: BY TARGET
; FILE REFERENCE: AVIGEN- CELLS
CURRENT APPLICATION NUMBER: 09-03743
; CURRENT FILING DATE: US/09/364, 862
; EARLIER APPLICATION NUMBER: 60/125, 974
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/104, 994
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
LENGTH: 11933

```

[illegible]

Db	901	gtccctgaagagatgtrgccaatgctctcgaccacatgctgccttaactaactatcaatct	960
Oy	961	tctcatgtggaacctggttaagaagcttaattcaagccotta	1000
Db	961	tctcatgtggaacctggttaagaagcttaattcaagccotta	1000

RESULT 3
 US-09-470-618-14
 Sequence 14, Application US/09470618
 Patent No. 6200560
 GENERAL INFORMATION:
 APPLICANT: Conto, Linda B.
 APPLICANT: Colosi, Peter C.
 TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
 TITLE OF INVENTION: by Target Cells
 FILE REFERENCE: Avigen -04082
 CURRENT APPLICATION NUMBER: US/09/470,618
 CURRENT FILING DATE: 1999-12-22
 EARLIER APPLICATION NUMBER: 09/364,862
 EARLIER FILING DATE: 1999-07-30
 EARLIER APPLICATION NUMBER: 60/125,974
 EARLIER FILING DATE: 1999-03-24
 EARLIER APPLICATION NUMBER: 60/104,994
 EARLIER FILING DATE: 1998-10-20
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 14
 LENGTH: 4999
 TYPE: DNA
 ORGANISM:
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 US-09-470-618-14

[illegible]

Db 470 agtgcacccgaagatcattcactcttggatggatggaaactgtcatatggagcattatgatcaagat 529
|||||
QY 541 gatctcggtgagctgtgctcttggacgcaagattccctctctagatggtccaaaacttttcca 600
530 gatctcggtgagctgtgctcttggacgcaagattccctctctagatggtccaaaacttttcca 589
QY 601 ttcaacacctcgaatcgtgtgtacaaaagaactctgtttgtgtgaattcaacgagatccatttc 660
590 ttcaacacctcgaatcgtgtgtacaaaagaactctgtttgtgtgaattcaacgagatccatttc 649
Db 661 aacatcgctaaagcgaagccacccttgatgtgtcgtctagttctccatccagcttga 720
650 aacatcgctaaagcgaagccacccttgatgtgtcgtctagttctccatccagcttga 709
QY 721 gttctatgatatacgtgtgtcattacacttaagaacatgtgcttccatcccttgcagcttcat 780
710 gttctatgatatacgtgtgtcattacacttaagaacatgtgcttccatcccttgcagcttcat 769
Db 710 gttctatgatatacgtgtgtcattacacttaagaacatgtgcttccatcccttgcagcttcat 769
QY 781 gctgtttgtgtatccttactctgaaagctctcttgaggaagcttgaataatgatgttaagaccat 840
770 gctgtttgtgtatccttactctgaaagctctcttgaggaagcttgaataatgatgttaagaccat 829
Db 841 caaagggagaagaagaatgataaagtcttcccttggatgaagcacaataatgatgttgcag 900
830 caaagggagaagaagaatgataaagtcttcccttggatgaagcacaataatgatgttgcag 889
QY 901 gtctctgaagaagaatgtgccaatgtgcttgccttgagccactgtgcttactactatatactt 960
890 gtctctgaagaagaatgtgccaatgtgcttgccttgagccactgtgcttactactatatactt 949
Db 961 tctcatgtgaagctgtgtaaagaacttgaattcaggacctca 1000
950 tctcatgtgaagctgtgtaaagaacttgaattcaggacctca 989

```

US-09-364-862-14
RESULT 4
US-09-364-862-14
Sequence 14, Application US/09364862
Patent No. 6221349
GENERAL INFORMATION:
APPLICANT: Couto, Linda B.
APPLICANT: Colosi, Peter C.
TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
TITLE OF INVENTION: BY TARGET
TITLE OF INVENTION: CELLS
FILE REFERENCE: AVIGEN-03743
CURRENT APPLICATION NUMBER: US/09/364,862
CURRENT FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: 60/125,974
EARLIER FILING DATE: 1999-03-24
EARLIER APPLICATION NUMBER: 60/104,994
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 4999
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-364-862-14

```

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	57.78;	Score 577.4;	DB 2;	Length 5035;
Best Local Similarity	99.88;	Pred. No. 1.2e-173;		
Matches 578; Conservative	0;	Mismatches		

	-	Indels	Gaps
422 ccaccatgcaaatagacatcctcc-	1;	0;	0;

OY	422	caacatcgcaatlagagctctccacctgcctctcttcgtgccttcttgatctcgttca	48
Db	30	CCACCATGGAAATAGAGCTCTCCACGTCTTCTTTGTGCTTGGATTCTGCTTAA	85
OY	482	gtgccaccagaagaaactacactggtgtgcagtgtgaacgtcgtacgtacatactgaagt	54
Db	90	GTGCCACCGAAGAACTACTCTGGGTGCAAGTGAACCTGCATGGGACTATATGCCAAGT	14
OY	542	atctgtgtgaaactgcctgtgtgaagcaagaattccctccatagatgcgaattcttcac	60
Db	150	ATCTGTGGAACTCTCCGTGACGCAAGATTTCTCCTAAGATGCCAAATCTTTTCAT	20
OY	602	tcaacacctcagtcgtgtcaaaaagacccgtttgtagaattccagatcacccttca	66
Db	210	TCAACACCTCAGTCTGTGTCAAAAAAGACTCTGTTGTGAATTCACGGATCACTTTTCA	26
OY	662	acatcgctgaagccaagaagccacccctgatatggtctcaggtctccaccaccagcttagg	72
Db	270	ACATCGCTAAGCCAAAGCCACCTGTGAGGTGCTGTAGTCTTCACATCCAGGCTGAGG	32
OY	722	tttatgatcacgtgtgtcattcaacttaagaacatgtgcttccatccctgcagcttcattg	78
Db	330	TTTATGATACAGTGCATTTACATTTAAGAAATGGCTTCCATCTCTCAAGTTCATG	38
OY	782	ctgttgatgtatccctaccaggaagctctcgaaggagctgaatatatgtatcaagccagtc	84
Db	390	CTGTGTGTATCTTACTGGAAGCTTCTGAGGACCTGGAATATGATGATGACCAAGTC	44
OY	842	aaaggagaaagaagatgtatnaagctctccctgtgtgaagccatacattgtcttcagag	90
Db	450	AAAGGGAAGAAAGATGATAAAGTCTTCCCGTGGGAAGCCATACATATGCTGGCAGG	50

Query Match 57.7%; Score 577.4; DB 3; Length 5035;
 Best Local Similarity 99.8%; Pred. No. 1,2e-173;
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 422 ccacatgcaaatagagacttccactgctcttcttctgctcttggatctgttca 481
DB 30 CCACATGGAATAGAGCTCTCCACCTGCTTCTTCTGCTTTTGGCATTCGCTTAA 89
QY 482 gtgcacacaaagatactactgtgagtgagtgagtgagtgagtgagtgagtgag 541
DB 90 GTGCCACCAAGATAGACTACTGCTGGGAGTGAGTACTGTCATGAGTATATGCAAGTG 149
QY 542 atctcggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 601
DB 150 ATCTCGGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 209
QY 602 tcaacacactgagtcgttcaaaaagactctgttggatgaattcagatcaaccttca 661
DB 210 TCAACACCTGAGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269
QY 662 acatgctgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 721
DB 270 ACATGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 329
QY 722 ttatgatacaatgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 781
DB 330 TTATGATACAGTGTGCTTACTTAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 389
QY 782 ctgttggtgtatctactgagaaagctcttgagagagagagagagagagagagag 841
DB 390 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 449
QY 842 aaagagagagagagagagagagagagagagagagagagagagagagagagagag 901
DB 450 AAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 509
QY 902 tccctgaagagagagagagagagagagagagagagagagagagagagagagag 961
DB 510 TCCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 569
QY 962 ctcatgtgactgtgtaaaagacttgatcagagcctca 1000
DB 570 CTCATGTGAGCTGTGTAAGAGCTTGAATTCAGGCTCA 608

```

RESULT 8
 5171844-1
 : Patent No. 5171844
 : APPLICANT: VAN OOVEN, ALBERT J.J.; PANNEKOEK, HANS; VERBET,
 : MARTINUS P.; VAN LEEN, ROBERT W.
 : TITLE OF INVENTION: PROTEINS WITH FACTOR VIII ACTIVITY
 : PROCESS FOR THEIR PREPARATION USING GENETICALLY-ENGINEERED CELLS
 : AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM
 : NUMBER OF SEQUENCES: 12
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/205,226
 : FILING DATE: 10-JUN-1998
 : SEQ ID NO: 1
 : LENGTH: 8241
 5171844-1

Query Match 57.5%; Score 575; DB 6; Length 8241;
 Best Local Similarity 100.0%; Pred. No. 9.2e-173;
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 426 catgcaaatagagcttccactgctcttcttctgctcttggatctgtctttagtc 485
DB 6 catgcaaatagagcttccactgctcttcttctgctcttggatctgtctttagtc 65
QY 486 caccagaagatactactgtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 545
DB 66 caccagaagatactactgtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 125

```

```

QY 546 cgtgagctgctgtgagcgaagatttccctcctcctcctcctcctcctcctcctcct 605
DB 126 cgtgagctgctgtgagcgaagatttccctcctcctcctcctcctcctcctcctcct 185
QY 606 cactcagctgtgtacaaaagactctgttltgaaatcagatcagatcagatcagatc 665
DB 186 cactcagctgtgtacaaaagactctgttltgaaatcagatcagatcagatcagatc 245
QY 666 cgttaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 725
DB 246 cgttaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 305
QY 726 tgatacagtgatattacacttaagaacatggtcttccatcctcctcctcctcctc 785
DB 306 tgatacagtgatattacacttaagaacatggtcttccatcctcctcctcctcctc 365
QY 786 tgggtatcctactggaagccttctgagggagctgaatataatgatacagaccagtc 845
DB 366 tgggtatcctactggaagccttctgagggagctgaatataatgatacagaccagtc 425
QY 846 ggaagaaagaatgataaagcttctcctcctcctcctcctcctcctcctcctcctc 905
DB 426 ggaagaaagaatgataaagcttctcctcctcctcctcctcctcctcctcctcctc 485
QY 906 gaaagaaatggttcaatgagcctcctgagccactgtgcttaactcaatcattctca 965
DB 486 gaaagaaatggttcaatgagcctcctgagccactgtgcttaactcaatcattctca 545
QY 966 tgtgacctgtgtaaaagacttgatcagagcctca 1000
DB 546 tgtgacctgtgtaaaagacttgatcagagcctca 580

```

RESULT 9
 US-08-366-851A-1
 : Sequence 1, Application US/08366851A
 : Patent No. 5681746
 : GENERAL INFORMATION:
 : APPLICANT: Bodner, Mordechai
 : APPLICANT: De Polo, Nicolas J.
 : APPLICANT: Hsu, David Chi-Tang
 : APPLICANT: Chang, Steven
 : TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII
 : NUMBER OF SEQUENCES: 3
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Viagene, Inc.
 : STREET: 11055 Roselle Street
 : CITY: San Diego
 : STATE: California
 : COUNTRY: U.S.A.
 : ZIP: 92121
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/366,851A
 : FILING DATE:
 : CLASSIFICATION: 514
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Chambers, Daniel M.
 : REGISTRATION NUMBER: 34,561
 : REFERENCE/DOCKET NUMBER: 930049.438
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (619) 452-1288
 : TELEFAX: (619) 452-2616
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 8967 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: both


```

;          TOPOLOGY: unknown
;          MOLECULE TYPE: CDNA
;          FEATURE:
;          NAME/KEY: CDS
;          LOCATION: 110..7165
;
US-08-366-851A-1

```

Query Match	57.5%;	Score 575;	DB 1;	Length 8967;
Best Local Similarity	100.0%;	Pred. No. 9.6e-173;		
Matches 575;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

OY	426	catgaaataagagctctccaccctgcctctctctctgcttttgatctcgtttagtc	485
Db	109	CATGCAATTAAGCTCTCCACCTGCTCTCTCTGCTTGGATTCGCTTAATGC	168
OY	486	caccagaatactactacctgaggtgcagtggaactctcatggaactatctgaagtgtct	545
Db	169	CACCAGAAGATACTAACCTGGGTGCAAGTGCATGGACATATATGCMAATGATCT	228
OY	546	cagtgagctgccctgtagcgcaagattctcctctagagtgccaatactttccatcaa	605
Db	229	CGGTAGCTGCTGCTGGACGCMAATTTCCTCCAGATGCGCAAAATCTTTCCATTCAA	288
OY	606	cacctcaatctgtgtaaaaaaagactcgtttgtagaatttaaggatlaacttttcaacat	665
Db	289	CACCTCACTGCTGTACAAAAAGACTCTGTTGTGTAAGATTACAGATACCTTTTCAACAT	348
OY	666	cgttaagccaagagccaccttgatctgaggtctctgtagagctcccatccagctgagtta	725
Db	349	CGCTAAGCAAGGCGCACCTGGATGGGTCTGCTAGTGCCTACATCCAGCTGAGGTTTA	408
OY	726	tgtataagttgtcatctaaccttaagaacatggtctcccatcctgtcagttcattatgctgt	785
Db	409	TGATACAGTGTCTCAATTCACATTAGAACAATGGCTCCCATCTGCTGAGTTCATGCTGT	468
OY	786	tgtgtatctccactcggagaagctctcttgaggagctcgtaatgatgatcgaagccagtaag	845
Db	469	TGTGCTATCTCTACTGGAAAGCTTCTTGAGGAGCTGGAATGATGATCAGCCAGTCAAG	528
OY	846	ggaagaagaagatbataaagtctctcccttgtagaagccacataatgtctcgaagttcct	905
Db	529	GGAGAAAGAAGATATTAAGTCTTCCCTGGTGGAAAGCCATACATATGTCTGGCAGGTCT	588
OY	906	gaaagaagaatgtctcaatgagcctctgagcccaatgtgccttaactactcaatattcttca	965
Db	589	GAAGAAGAATGTCTCAATGGCTCTTGACCCACATGTGCTTACTACTACATATCTTCTCA	648
OY	966	tgtgacctgtgtaaaagacttgatcattcagctcca	1000
Db	649	TGTGACCTGTGTAAGACTTTGAATTCAGGCCCTCA	683

RESULT 10
 US-07-864-004B-3
 : Sequence 3, Application US/07864004B
 : Patent No. 5364771
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Lollar, John S.
 : APPLICANT: Runge, Marshall S.
 : TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
 : NUMBER OF SEQUENCES: 6
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Kilpatrick & Cody
 : STREET: 1100 Peachtree Street
 : City: Atlanta
 : STATE: Georgia
 : COUNTRY: US
 : ZIP: 30309
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS

```

1      SOFTWARE: Patentin Release #1.0, Version #1.25
2      CURRENT APPLICATION DATA:
3      APPLICATION NUMBER: US/07/864,004B
4      FILING DATE: 07 APRIL 1992
5      CLASSIFICATION: 435
6      ATTORNEY/AGENT INFORMATION:
7      NAME: Padst, Patrea L.
8      REGISTRATION NUMBER: 31,284
9      TELECOMMUNICATION INFORMATION:
10     TELEPHONE: 404-815-6508
11     TELEFAX: 404-815-6555
12     INFORMATION FOR SEQ. ID NO.: 3:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 9009 base pairs
15     TYPE: nucleic acid
16     STRANDEDNESS: single
17     TOPOLOGY: linear
18     MOLECULE TYPE: cDNA
19     HYPOTHETICAL: NO
20     ANTI-SENSE: NO
21     FRAGMENT TYPE: N-terminal
22     ORIGINAL SOURCE:
23     ORGANISM: Homo sapien
24     TISSUE TYPE: Liver
25     FEATURE:
26     NAME/KEY: misc_feature (Domain Structure)
27     LOCATION: 5001 .. 7053
28     OTHER INFORMATION: /note="Equivalent to the A3-C1-C2"
29     OTHER INFORMATION: domain"
30     FEATURE:
31     NAME/KEY: misc_feature (Domain Structure)
32     LOCATION: 1 .. 2277
33     OTHER INFORMATION: /note="Equivalent to the A1-A2"
34     OTHER INFORMATION: domain"
35     US-07-864-004B-3

```

Query Match	57.3%	Score 573.4	DB 1:	Length 9009;
Best Local Similarity	99.8%	Pred. NO. 3.1e-172;		
Matches 574;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY 426	catgcaaatagagctctccacctgctctctctctgtgctcttttgatctcttgcatctcgtttagttgc	485		
Db 150	CATGCAAAATAGAGCTCTCCACCTCGCTCTTCTTGCGCTTTGGCATTCGCTTAGTGC	209		
QY 486	caccagaagatatactaccctcgggtgtgcaggttggaactgtgcattgaggactctatgcaaaagtgacct	545		
Db 210	CACCAAGAAATACTACCTCGGGTGGACGTGGAACTGATCGAGGACATATATGCAAAAGTGACTT	269		
QY 546	cagttgagctgtgctcttgcagcgcaagaatcttcctcctctagatgtgccaaaatctttccatcaa	605		
Db 270	CGGTAGCTGCTGCTGGACGCAAGATTTCCGTCGTAGAGTGCAGAAATCTTTTCCATTCAA	329		
QY 606	cacctcagtcctgtatcaaaaaagactctgtttgtgaaatcaacgatacccttttcaaat	665		
Db 330	CACCTCATCTCTGTACAAAAGACTCTGTTTGTATTAATTCACGGTTCCACCTTTTCAACAT	389		
QY 666	cgttaagcaagaagcacccttgatgtgctctgtagtgcctacatccacacagcgttgaagtta	725		
Db 390	CGCTAAGCCAAAGGCCACCCTGGATGGGCTCTAGGCTCTACCATCCAGGCGTAGGTTTA	449		
QY 726	tgtatacagttgttcatttaactttaagaacaatgtgcttccatccctgtcagttcttcatgtctgt	785		
Db 450	TGATATCACTGGTCATTTACATTTAAGAAATATGGCTTCCATCTGTCACTCTTTCATAGCTGT	509		
QY 786	tgtgtatccctaccaggaaagctctcgaagagagctgaaatatagatatacagacagtgcaag	845		
Db 510	TGGTGTATCTACTACGGAAAGCTTCTGAGGAGGAGCGTAATATATGATGATCGACACAGTCAAG	569		
QY 846	ggagaaagaagaatataaagctctccctctgtgtgaaagccatacatatgctctggcaggttctt	905		
Db 570	GGAGAAAGAATATATTAAGTCTTCCCTGGTGGAAAGCATTCATATGTCCTGGCAGGTCCT	629		

QY 906 gaaagaatggtcaatggtcctgagccactgctgactactactactactcctca 965
 Db 630 GAAAGAGATGGTCCATGAGCTGACCCACTGAGCTTACTACTACTACTACTCTCA 689
 QY 966 tftgagacctgttaagaacttgaattcagagctca 1000
 Db 690 TGTGACCTGTGTAAGACTTGATTCAGGCTCA 724

RESULT 11

US-08-251-937A-3
 Sequence 3, Application US/08251937A
 Patent No. 5583209
 GENERAL INFORMATION:
 APPLICANT: Lollar, John S.
 APPLICANT: Runge, Marschall S.
 TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kilpatrick & Cody
 STREET: 1100 Peachtree Street
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: US
 ZIP: 30309
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/251,937A
 FILING DATE: 31-MAY-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/864,004
 FILING DATE: 07-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Pratt, John S.
 REGISTRATION NUMBER: 29,476
 REFERENCE/DOCKET NUMBER: EMU106DIV
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 404-815-6367
 TELEFAX: 404-815-6555
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9009 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 ORIGINAL SOURCE:
 ORGANISM: Homo sapien
 TISSUE TYPE: Liver
 FEATURE:
 NAME/KEY: misc_feature (Domain Structure)
 LOCATION: 5001 . . . 7053
 OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2
 OTHER INFORMATION: domain"
 NAME/KEY: misc_feature (Domain Structure)
 LOCATION: 1 . . . 2277
 OTHER INFORMATION: /note= "Equivalent to the A1-A2
 OTHER INFORMATION: domain"
 US-08-251-937A-3

Query Match 57.3%; Score 573.4; DB 1; Length 9009;
 Best local Similarity 99.8%; Pred. No. 3,1e-172;

Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 426 catgcaataagatcttcacactctctctctctctctctctctctctctctctc 485
 Db 150 CATGCAATAAGATCTTCACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 209
 QY 486 caccagaataactactacactcgtggtcagtggaactgltcattggaactatata 545
 Db 210 CACCAGAATAACTACTACCTGGGTGCAGTGAACCTGTCAATGGAGATATGCA 269
 QY 546 cgtgtgagctgctctgtagcgcgaagaattctctcctcgaagtgccaaatctt 605
 Db 270 CGGTGAGCTGCTGCTGGAGCGCAAGATTCTCTAGAGTGCCAAATCTTTCCATT 329
 QY 606 caccctacgtctgtacaaaagaactctgtttagaattcagagtaacctttca 665
 Db 330 CACCTCAGTCTGTACAAAAGACTCTGTGTGAAATTCAGGTTCACTTTCAACAT 389
 QY 666 cgttaagccaagggccaccctgtggtggtcgtgtagtctcaacatcagagct 725
 Db 390 CGCTAAGCCAAAGCCACCCCTGATGGGTCTGCTAGTCTTCCATCCAGGCTGA 449
 QY 726 tgatacagtggtcattactaactaagaatggtctccatctctgagcttca 785
 Db 450 TGATACAGTGTGCTATACACTTAAGAACATGCTTCCATCTCTGATGCTTCA 509
 QY 786 tgggtatcctactctgaaagctctcgaagggagctgaatatgatcagaagca 845
 Db 510 TGGTATCTCTACTTGAAAGCTTCTGAGGAGCTGAATATGATATCAGACCAAT 569
 QY 846 ggaagaagaatgataagttctccctgtgtggaagccatacatatgltcgt 905
 Db 570 GGAGAAAGATGATGATAAGTCTTCCCTGTGGAAGCCATGATGCTTGCGAG 629
 QY 906 gaaagaatggtcgaatggtcctcgaaccactggtcttaccactacatctt 965
 Db 630 GAAAGAGATGGTCCAAATGGCTTGAGCCACTGAGCTTACTACTACTACTCT 689
 QY 966 tftgagacctgttaagaacttgaattcagagctca 1000
 Db 690 TGTGACCTGTGTAAGACTTGATTCAGGCTCA 724

RESULT 12

US-08-212-133A-1
 Sequence 1, Application US/08212133A
 Patent No. 5663060
 GENERAL INFORMATION:
 APPLICANT: Lollar, John S.
 APPLICANT: Runge, Marschall S.
 TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kilpatrick & Cody
 STREET: 100 Peachtree Street
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: US
 ZIP: 30303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/212,133A
 FILING DATE: March 11, 1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/864,004
 FILING DATE: 07-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Prabst, Patricia L.

01 966 tbtgaccctgttaaagacttbaatcagcccta 1000
02 |||||||
03 |||||||
04 |||||||
05 |||||||
06 |||||||
07 |||||||
08 |||||||
09 |||||||
10 |||||||
11 |||||||
12 |||||||
13 |||||||
14 |||||||
15 |||||||
16 |||||||
17 |||||||
18 |||||||
19 |||||||
20 |||||||
21 |||||||
22 |||||||
23 |||||||
24 |||||||
25 |||||||
26 |||||||
27 |||||||
28 |||||||
29 |||||||
30 |||||||
31 |||||||
32 |||||||
33 |||||||
34 |||||||
35 |||||||
36 |||||||
37 |||||||
38 |||||||
39 |||||||
40 |||||||
41 |||||||
42 |||||||
43 |||||||
44 |||||||
45 |||||||
46 |||||||
47 |||||||
48 |||||||
49 |||||||
50 |||||||
51 |||||||
52 |||||||
53 |||||||
54 |||||||
55 |||||||
56 |||||||
57 |||||||
58 |||||||
59 |||||||
60 |||||||
61 |||||||
62 |||||||
63 |||||||
64 |||||||
65 |||||||
66 |||||||
67 |||||||
68 |||||||
69 |||||||
70 |||||||
71 |||||||
72 |||||||
73 |||||||
74 |||||||
75 |||||||
76 |||||||
77 |||||||
78 |||||||
79 |||||||
80 |||||||
81 |||||||
82 |||||||
83 |||||||
84 |||||||
85 |||||||
86 |||||||
87 |||||||
88 |||||||
89 |||||||
90 |||||||
91 |||||||
92 |||||||
93 |||||||
94 |||||||
95 |||||||
96 |||||||
97 |||||||
98 |||||||
99 |||||||
00 |||||||

Db 690 tbtgaccctgttaaagacttbaatcagcccta 724

RESULT 13
US-08-474-503-1
Sequence 1, Application US/08474503
Patent No. 5744446
GENERAL INFORMATION:
APPLICANT: Emory University
TITLE OF INVENTION: Hybrid Human/Animal Factor VII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,503
FILING DATE: 07-JUN-1995
CLASSIFICATION: 433
ATTORNEY/AGENT INFORMATION:
NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EM0106CIP(3)
COMMUNICATION INFORMATION:

```

1 TELETYPE: FOR SEQ. ID NO.: 1:
2 INFORMATION CHARACTERISTICS:
3 SEQUENCE CHARACTERISTICS:
4 LENGTH: 9009 base pairs
5 TYPE: nucleic acid
6 STRANDEDNESS: single
7 TOPOLOGY: linear
8 MOLECULE TYPE: cDNA
9 HYPOTHETICAL: NO
10 ANTI-SENSE: NO
11 ORIGINAL SOURCE:
12 ORGANISM: Homo sapien
13 TISSUE TYPE: Liver
14 FEATURE:
15 NAME/KEY: misc_feature (Domain Structure)
16 LOCATION: 5125..7053
17 OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2
18 OTHER INFORMATION: domain"
19 FEATURE:
20 NAME/KEY: misc_feature (Domain Structure)
21 LOCATION: 1..2277
22 OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain."
23 FEATURE:
24 NAME/KEY: Domain
25 LOCATION: 1..2277
26 OTHER INFORMATION: /note= "cDNA encoding human factor
27 OTHER INFORMATION: VIII."
28
29 US-08-474-1

```

	Query Match	57.38;	Score 573.4;	DB 1;	Length 9009;	
	Best Local Similarity	99.88;	Pred No. 3.1e-172;			
	Matches 574; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;	
OY	426 catgcaaatagagctctccacacctcctcttccttgcgtcgatcttgcatcttaatgc					485
Dd	150 CATGCANATNAGCTCTCCACCTGCTTTTCGTGGCTTTTGCAATTCTGTTTTATGTC					209

RESULT 14
 US-08-670-707A-1
 Sequence 1, Application US/08670707A
 Patent No. 5859204
 GENERAL INFORMATION:
 APPLICANT: Lollar, John S
 TITLE OF INVENTION: Hybrid Human/Animal Factor VII
 NUMBER OF SEQUENCES: 140
 CORRESPONDENCE ADDRESS:
 ADDRESS: Greenlee, Winner and Sullivan, P.C.
 STREET: 3370 Manhattan Circle Suite 201
 CITY: Boulder
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/670,707A
 FILING DATE: 26-JUN-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US94/13200
 FILING DATE: 15-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/212,133
 FILING DATE: 11-MAR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/864,004
 FILING DATE: 07-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Greenlee, Lorraine L.

```

1  REGISTRATION NUMBER: 27,984
2  REFERENCE/DOCKET NUMBER: 75-95F
3  TELECOMMUNICATION INFORMATION:
4  TELEPHONE: 303/499-8080
5  TELEFAX: 303/499-8089
6  INFORMATION FOR SEQ ID NO: 1:
7  SEQUENCE CHARACTERISTICS:
8  LENGTH: 9009 base pairs
9  TYPE: nucleic acid
10  STRANDEDNESS: double
11  TOPOLOGY: not relevant
12  MOLECULE TYPE: cDNA to mRNA
13  HYPOTHETICAL: NO
14  ANTI-SENSE: NO
15  ORIGINAL SOURCE:
16  ORGANISM: Homo sapiens
17  TISSUE TYPE: Liver
18  FEATURE:
19  NAME/KEY: misc-feature
20  LOCATION: 5125..7053
21  OTHER INFORMATION: /product= "Domain Structure"
22  OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2 domain"
23  FEATURE:
24  NAME/KEY: misc-feature
25  LOCATION: 1..2277
26  OTHER INFORMATION: /product= "Domain Structure"
27  OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
28  FEATURE:
29  NAME/KEY: misc-feature
30  LOCATION: 1..2277
31  OTHER INFORMATION: /product= "Domain"
32  OTHER INFORMATION: /note= "cDNA encoding human factorVIII"
33  US-08-670-707A-1

```

Query	Match	57.3%	Score 573.4	DB 2	Length 9009
Best Local Similarity	99.8%	Pred. No. 3.1e-172			
Matches 574	Conservative 0	Mismatches 1	Indels 0	Gaps 0	
QY	426	catgcaaatagagctctccacctgctctctctctgctgcttttgagcttcgctttagtc	485		
Db	150	CATGCAAAATAGAGCTCCACCTGCTCTTCTTGCTGCTTTGGCAATTCGCTTAGTGC	209		
QY	486	caccagaagaataactactctggtggtcagttgaaactgtcaatggaactaatgcaagtgact	545		
Db	210	CACCAGAAGATACTACTCTGGTGCAGTGAACGTCACTGAGCATATATGCCAAGTGAATCT	269		
QY	546	cgggtgagctctgcttggacgcaagacttctcctctgaaagtgcgaatacttttccatcaa	605		
Db	270	CGGTGAGCTCTGCTGTGAGCGCAAGATTTCTCTTACATGTCGCAAAATCTTTTCATTCAA	329		
QY	606	caacctagtcgtgtacaaaagaactctgttttgtaaatccagatccacttccaacat	665		
Db	330	CACCTCAGTGTGTNAAAAACATCTGTTGTAAATTCAGGATTCACCTTTCAACAT	389		
QY	666	cggtaagcaaaagccacccctggaatgggtctcgtagtcctccaccctcaggctgaagttta	725		
Db	390	CGGTAGCCCAAGGCCACCCCTGGATGGGTCCTCTAGTCCTACCATCCAGGCTAAGGTTTA	449		
QY	726	tgatacagttgycatttacccttaagaaoatggcttcccacccctgtcaagcttcacgtct	785		
Db	450	TGATCAGATGTGTCATTTACATTAAAGAACATAGGCTCCCATCTGTCAAGTTCATGCTGT	509		
QY	786	tgtgtatccctaacgtgaaagcttcgtgaggtctgaaatgtaatgaaatcaaacgtccaag	845		
Db	510	TGTGTATCCCTAACGTGAAAGCTTCTGTAGGAGCTGAATATGATGATCAACCAAGTCAAG	569		
QY	846	ggagaagaagaatataaagtccttcctctggtggaaagcaatacatatgtctgcaagtcct	905		
Db	570	GGAGAAGAAGATGATTAAGTCTTCCCTGCTGTAAGCCATACATATGTCGTGCAAGTCTCT	629		
QY	906	gaaaggaatagttccaaatgagcctctgaaccacactgtgcttactactatatacttcttca	965		

Db 630 GAAAGAGATGTCACATGAGCCTCTGACCCAGCTGCTTACTACTATCTTTCTCA 689
 Qy 966 tttgacctgttaaaagacttaattcaagcctca 1000
 Db 690 TGTGACCTGTAAAGACTTGAAATTCAGGCTCA 724

RESULT 15
 US-09-037-601-1
 ; Sequence 1, Application US/09037601
 ; Patent No. 6180371
 ; GENERAL INFORMATION:
 ; APPLICANT: Lollar, John S.
 ; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Wanner and Sullivan, P.C.
 ; STREET: 5370 Mannheim Circle Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/037,601
 ; FILING DATE: 26-JUN-1996
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/US94/13200
 ; FILING DATE: 15-NOV-1994
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/212,133
 ; FILING DATE: 11-MAR-1994
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/864,004
 ; FILING DATE: 07-APR-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feiber, Donna M.
 ; REGISTRATION NUMBER: 33,878
 ; REFERENCE/DOCKET NUMBER: 75-95F
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 303/499-8080
 ; TELEFAX: 303/499-8089
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9009 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: cDNA to mRNA
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; TISSUE TYPE: Liver
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 5125..7053
 ; OTHER INFORMATION: /product= "Domain Structure"
 ; OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2 domain"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..2277
 ; OTHER INFORMATION: /product= "Domain Structure"
 ; OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..2277
 ; OTHER INFORMATION: /product= "Domain"
 ; OTHER INFORMATION: /product= "Domain"

OTHER INFORMATION: /note= "CDNA encoding human factorVIII"
 US-09-037-601-1

Query Match 57.3%; Score 573.4; DB 4; Length 9009;
 Best Local Similarity 99.8%; Pred. No. 3.1e-172;
 Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 426 catgcaaatagagctctccaccctgctctttctgtgctcttttcgacttcgtcttagtgc 485
 Db 150 CATGCAAAATAGAGCTCTCCACCTGCTTCTTGTGCTTTTCGATTCGTTTGTGTC 209
 Qy 486 caccagaagaatactacactgtgtgcaatgtaactgttaactggaatataagaaatgact 545
 Db 210 CACCAGAAGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269
 Qy 546 cgtgtaagctgctgtggaagcaagattctctctagagtgagtgcaaaatctttccattca 605
 Db 270 CGGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 329
 Qy 606 caccctagctgtgtgcaaaagactctgttctgagaatcaagatcaacctttcaaat 665
 Db 330 CACCTCAGCTGTGTACAAAAAGACTCTGTTGTAGATTCACGGTTTCACTTTCAACAT 389
 Qy 666 cgtcaagccaagcaccctggaatggtgtgctgtaagtcctaccatccaggctgaagtta 725
 Db 390 CGCTAAGCAAGGCCACCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 449
 Qy 726 tgatacagtggtcaatcaacttaagaacatggtctccatccctgtaagtcctcaatgctgt 785
 Db 450 TGATACAGTGTGTCACTTACACTTAAGAACATGCTTCCATCTCTGCTGCTGCTGCTGCTGCT 509
 Qy 786 tgggtatccctaccggaagctctggaaggtggaatgataatgataatgataatgataatg 845
 Db 510 TGGTGTATCTTACTGGAAGCTTCTGAGGAGCTGAATATGATGATGATGATGATGATGATGAT 569
 Qy 846 ggaagaagaatgataaagctctccctggtgtggaagcacaataatgctggaagctct 905
 Db 570 GGAGAAAGAAATGATTAAGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 629
 Qy 906 gaaagagatggtccaatggtcctggaaccactgtgcttaccctaccataatcttctca 965
 Db 630 GAAAGAGATGTCACATGAGCCTCTGACCCAGCTGTGCTTACCTACTACTATCTTTCTCA 689
 Qy 966 tttgacctgttaaaagacttaattcaagcctca 1000
 Db 690 TGTGACCTGTAAAGACTTGAAATTCAGGCTCA 724

Search completed: January 17, 2002, 16:26:40
 Job time: 12310 sec

Tue Jan 22 15:12:28 2002

us-09-740-211-13_copy_1_1000.rni

Page 12

VERSION	BF150467.1	GI:11031862
KEYWORDS	EST.	
SOURCE	house mouse	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 669)	
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-riemail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov/image/html/lresources.shtml	
FEATURES	source	
	MG1:1427189 Seq primer: -40RP from GIDCO High quality sequence stop: 428. Location/Qualifiers 1..669 /organism="Mus musculus" /strain="C57/B6" /db_xref="taxon:10090" /clone="IMAGE:3666421" /clone_1lb="NCI-CGAP_Mam5" /tissue_type="tumor, gross tissue" /dev_stage="7 months" /lab_host="DHI0B" /note="Organ: mammary; Vector: PCMV-SPOB6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"	
BASE COUNT	213 a 126 c 142 g 183 t 5 others	
ORIGIN		
Query Match	9.3%; Score 92.6; DB 11; Length 669;	
Best Local Similarity	54.2%; Pred. NO. 2.4e-16;	
Matches 207; Conservative	0; Mismatches 172; Indels 3; Gaps 1.	
OY	619 tacaanaagacitgttltagaatcaaggaatcacccttcaacatcgtctaaagcaagg 678	
Db	211 TATAGAAAGGCCCTTATTTTTTCAGTACACAGATGGCAGCTTATAGTAACTATAGCAAA 270	
OY	679 caaccctgatgtgctcgtactgaagctcactacatccaggtgaagttatgatcacgtgtgc 738	
Db	271 CCAGCTGTGGCTGTAGGGCTTTTAGGCCCTGTCAACAAAGCTGAAGTTGAAGATTAAGCTTAT 330	
OY	739 attaaccttaagaacctgtgtctccactcctgtcgaagctcctcaagctgtgtgtatcctac 798	
Db	331 GTTCACCTTAAAGAACCTTCCTCTAGGATCTTACACTTTTCAATCAGACATGGGTAACGTAC 390	
OY	799 tggaaagctcttgaaggagatgtaattgtatgtatcagacagtcacagtcacaaagggagaagaagt 858	
Db	391 ACCAAGAGATATGAGGAGGCGGCTACCTGACCAACACCACTGATTTTCAACAGGCGCTGAT 450	
OY	859 gataaagctctccctgtgtggaagccaatacatatgtctgtgcaagttcctgtaagaagaaatgt 918	
Db	451 GACAAGATGCTTCCCGGACAAACAGTATGTGTATGTCTGATGCTGATGCC---AATGAGCCAAAG 507	
OY	919 ccaatgacctctgaccacacgtgtccttaacctactcatatcttccatgtgagacctgtga 978	
Db	508 CCTGGAGAGGAGACACACATTTGTGTGACCAAGATTTTACACATGCCAATGTTATGCTGAC 567	
OY	979 aagaacttgaaatcagagcctca 1000	

Db	568	AAAGATATTCAGATCAGACATCA	589
RESULT	4		
LOCUS	AA461838		
DEFINITION	AA461838 477 bp mRNA	EST	10-JUN-1997
ACCESSION	AA461838		
VERSION	AA461838.1	GI:2186729	
KEYWORDS			
SOURCE	mus musculus.		
ORGANISM	mus musculus.		
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus. 1 (bases 1 to 477)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellander,B., K., Stepien,M., Tan,F., Underwood,R., Moore,B., Thelning,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouse@wustl.wustl.edu This clone is available royalty-free through LNC ; contact the IMAGE Consortium (info@image.lln.gov) for further information. MGI:503679 Seq primer: -28ml3 rev2 ET from Amersham High quality sequence stop: 413.		
FEATURES			
Source	Location/Qualifiers		
	1..477		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:851527"		
	/clone_id="Soares_mammary_gland_NbMNG"		
	/sex="male"		
	/tissue_type="mammary gland"		
	/dev_stage="4 weeks"		
	/lab_host="DH10B"		
	/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5',		
	TGTTACCAATCTGAAGTGGGCGGCGCGAAGCTTTTTTTTTTTTTTTTTTTT		
	T3]; double-stranded cDNA was ligated to Eco RI		
	adaptors (Pharmacia), digested with Not I and cloned into		
	the Not I and Eco RI sites of the modified pT73 vector.		
	RNA provided by Dr. Minoru Ko, Wayne State Univ. library		
	constructed and normalized by Bento Soares and M.Fatima		
	Bonaldi."		
BASE COUNT	148 a 102 c 102 g 125 t		
ORIGIN			
Query Match	9.2%; Score 92.4; DB 10; Length 477;		
Best Local Similarity	54.5%; Pred. No. 2.4e-16;		
Matches	208; Conservative 0; Mismatches 171; Indels 3; Gaps 1;		
QY	619 tacaaaagactctgttttagaatcaaggtacaccttttaacatcgttaagccaag 678		
DB	77 TATTAAGAGCCCTTTATTTTGAATACACAGATGGACCTTTAGTAGACTTTAGACAA 136		
QY	679 ccaccctgatggctcgtcgtaccacacacagcaggtcaggtatattatcagtgctc 738		
DB	137 CCAGCCTGGGTAGGTTTTTAGCCCTGTCATCAAGCTGAGGTTGAAGATTAAGTTAT 196		

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 521)	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	Unpublished (1997)
2	Emmert-Buck, M.D., Ph.D.	CDNA Library Preparation: M. Bento Soares, Ph.D.	CDNA Sequencing by: Washington University Genome Sequencing Center	Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.lnl.liv.gov/bbrp/image/image.html
3	Insert Length: 656	Std Error: 0.00	Seq primer: -40m13 fwd. Err from Amersham	High quality sequence stop: 360.
4	Location/Qualifiers	1. 521	/organism="Homo sapiens"	/db_xref="taxon:9606"
5	/clone="IMAGE:1455379"	/clone_1ib="NCI-CGAP_K1d5"	/tissue_type="2 pooled tumors (clear cell type)"	/lab_host="DH10B"
6	/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', AACGGAAGAATTCGGCCGCGCAATATATTTTATTTTATTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."			
7	BASE COUNT	138 a	100 c	107 g
8	ORIGIN	176 t		
9	Query Match	9.1%	Score 91.2:	DB 10; Length 521;
10	Best Local Similarity	52.3%	Pred. No. 5.7e-16;	
11	Matches 201; Conservative	0;	Mismatches 183;	Indels 0; Gaps 0;
12	617	tgatacaaaagactcgttctgtagaattcaagatcacccttcaacaacgctaagcca	676	
13	Db	425	TATATTAAGAAAGCCCTTATCTCAGTACACAGATGAAGCAACTTAGGCACTATAGAAA	366
14	QY	677	ggccaccctgatgatgctgctgtagtccctaccaccgagctggaaggttatgatacagtgg	736
15	Db	365	AACCGGTGCGCTTGGGTTTATGAGCCCTATATTCAAGCTGAAGCTGGAATTTAAAGTTT	306
16	QY	737	tcattacaacttaagaacatgctccaccatccctgcagtcctcatcagctgctgtgtatcct	796
17	Db	305	ATGTACACTTAAAAAACCTTGCTCTGAGGCCCTACACCTTCACTTCCATGCAATTA	246
18	QY	797	actggaagacttcgagggagctgaatatgatcatcagcagcagcaagggagaagaag	856
19	Db	245	ACTATTAAGGAACATGAGGGGGCCATTCACCTGTGTAACACCAAGATTTTCAAGAGCAG	186
20	QY	857	atgataagcttcctccctggtggaagcacaatacatgtctgacagctctgaagaagatg	916
21	Db	185	ATGACAAAGTATATCCAGAGAGACGATATACATATCATGTGCTTCCACATGAAACAAA	126
22	QY	917	gtccaatgagcctctgaccacatgctccatcactactacatattcttccatgltgagctgg	976
23	Db	125	GTCTCTGGGGAAGGAGATGTCATGTGTGACTAGATTTTACATTCACATTCATGATGCTC	66
24	QY	977	taaaagactgaattcaagcctca	1000
25				

Db	65	CAAAAGATATTGCCTCAGCACTCA	42
RESULT	12		
LOCUS	AL556703	918 bp	EST
DEFINITION	AL556703 LTR_NFL006.Pl2	Homo sapiens cDNA clone	CSODK005Y102 5
ACCESSION	AL556703	prime, mRNA sequence.	
VERSION	AL556703.1	GI:12899623	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 918)		
TITLE	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
FEATURES	Genoscope		
Source	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seque@genoscope.cns.fr Web : www.genoscope.cns.fr.		
	Location/Qualifiers		
	1..918		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CSODK005Y102"		
	/clone.lib="LTR_NFL006.Pl2"		
	/tissue.type="placenta"		
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	303 a	169 c	192 g
ORIGIN			253 t
			1 others
Query Match	9.1%	Score 91.2	DB 10; Length 918;
Best Local Similarity	52.3%	Pred. No. 6.9e-16;	
Matches 201; Conservative	0;	Mismatches 183;	Indels 0; Gaps 0;
QY	617	tgtaaaaaagctctgttctgtatgaatcaagatcaaccttcaacatcgtaagccaa	676
Db	199	TATATAAGAGAGGCCCTTATCTTACGTACACAGATGAAACCTTTAGGACAACTATAGAA	258
QY	677	ggccaccttgatggtctctgtaagtcctacatccagcgtgaagttatgatacagtyg	736
Db	259	AACCGGTCTGGCTTGCGTTTAAAGCCCTATATATCAAGAGCTGAAAGTGAAGATT	318
QY	737	tcaatcaactaagaagctctgtccatccatccctgtgaagcttcatgctgttgatcct	796
Db	319	ATGTACTACTTAAAAAACCCTTGCTTACAGCCCTTACACCTTTTACATGGAATTAACCT	378
QY	797	actggaagctctctgagagagctgaataatgataatcagcagctcaagaagaggaagaag	856
Db	379	ACTATAAGAGACATGATGAGGGGCACTACCTCGATTAACACACACACAGATTTC	438
QY	857	atgataaagctctccctctgtygtygaagccatataatgctctgagtgacctgaagaagatg	916
Db	439	ATGACAAAGTATATCCAGAGAGCAGTATACATACATGTTCTTGCCACTAGAGAACAA	498
QY	917	gtccaatggtccttggaaccactgtgaccttaactaactaacttcttcaatgtgaacctgg	976
Db	499	GTCTCTGGGAGAGAGATGCGCAATTGTGTGACTAGAGATTACCAATTCACCAATTAATGCTC	558
QY	977	taaaagacttgaaatcagagcctca	1000

Db 559 CAAAAGATATTCCTCAGACTCA 582

||||| |

RESULT 13

LOCUS BG862595 632 bp mRNA EST 29-MAY-2001

DEFINITION 602795978F1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4917258 5', mRNA sequence.

ACCESSION BG862595

VERSION BG862595.1 GI:14213133

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1. (bases 1 to 632)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: L1AM10828 row: k column: 19
High quality sequence stop: 631.
Location/Qualifiers

FEATURES

source

1. .632

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:4917258"

/clone_lib="NIH_CGAP_Mam4"

/tissue_type="tumor, gross tissue"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert 2.5 kb. Library constructed by Life Technologies, catalog # 12018-016. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP Library."

BASE COUNT 195 a 125 c 139 g 173 t

ORIGIN

Query Match 9.1%; Score 90.8; DB 11; Length 632;
Best Local Similarity 54.2%; Pred. No. 7.9e-16;
Matches 207; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

619 tacaaagaactctgtgtgtagaatcaagcaccctttcaacatcgtcaagccaaag 678
||||| |

Db 210 TATTAAGAAGGCCCTTATTTGAGTACACAGATGGCACCTTAGTCAGACTATAGACAAA 269
||||| |

Qy 679 gcaacctgagtggtctgctagctccatccagctcaggtggttaagatagatggtc 738
||||| |

Db 270 CCAAGCTTGCTAGGGCTTTTAAAGCCCTGTGCATCAAGCTGAAGTTGAAGATTAAGCTTAT 329
||||| |

Qy 739 attaacactaagaacatggtctccatccctgctcagcttcctcatgctgtgtgtatccctac 798
||||| |

Db 330 GTTCACTTAAAGAACCTTGCTCTAGATCTACACTTTTCATGACATGGGGTAAACGTAC 389
||||| |

Qy 799 tggaaagcttcgagggagctgaatatatgatacagaccagctcaagggagaaagaagat 858
||||| |

Db 390 ACCAAGAGATATGAGGAGCCCTCTACCTTGACACACACACTGATTTTCAACGGGCTGAT 449
||||| |

Qy 859 gataaagcttcctccgtggaagcacaataatgctgacaggtcctgaaagaagatggt 918
||||| |

Db 450 GACAAAGCTTCCCGACACAGATGTGTATGTCGTCATGCC---AATGAGCCAAGT 506
||||| |

Qy 919 ccaatgctctgacccactgctgcttactactcatatcttctcatgtagctgtgta 978
||||| |

Db 507 CCTGGAGAGGAGACAGCAATGTGTGACCAAGGATTTACCACTCCCATGTTATGCTCCA 566
||||| |

Qy 979 aaagacttgatcagagctca 1000
||||| |

Db 567 AAGATATTTGCATCAGACTCA 588
||||| |

RESULT 14

LOCUS AV653260 634 bp mRNA EST 07-SEP-2000

DEFINITION AV653260 GLC Homo sapiens cDNA clone GLCDBA09 3', mRNA sequence.

ACCESSION AV653260

VERSION AV653260.1 GI:9874274

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1. (bases 1 to 634)

AUTHORS Qian, B., Wu, T., Huang, Q., Huang, C., Kang, B., Gao, X., Xu, Z., Xiao, H., Xu, X., Li, N., Peng, Y., Liu, F., Ou, D., Song, H., Cheng, Z., Ou, J., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gu, Y., Chen, Z., and Han, Z.
Homo sapiens cDNA clone
Unpublished (2000)

TITLE Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

JOURNAL

COMMENT

FEATURES

source

1. .634

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="GLCDBA09"

/clone_lib="GLC"

/tissue_type="corresponding non cancerous liver tissue"

/dev_stage="Adult"

/lab_host="SOLR"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 213 a 119 c 128 g 172 t 2 others

ORIGIN

Query Match 9.0%; Score 89.6; DB 10; Length 634;
Best Local Similarity 52.1%; Pred. No. 1.8e-15;
Matches 200; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

617 tgtacaaaagaactctgtgtgtagaatcagagataccctttcaacatcgtcaagccaa 676
||||| |

Db 203 TATTAAGAAGGCCCTTATTTCTCAGTACACAGATGAACCTTTAGGACAACTATAGAAA 262
||||| |

Qy 677 ggcaccctgagtggtctgctagctccatccatccagctcaggtggttaagatagatggtc 736
||||| |

Db 263 AACCGCTTGCTAGGGCTTTTAAAGCCCTTATTAACAAGCTGAAGCTGAAGATAAAGTTT 322
||||| |

Qy 737 tcaataactaagaacatggtctccatccctgctcagcttcctcatgctgtgtgtatccct 796
||||| |

Db 323 ATGTACACTTAAATAAAACCTTGCTCTAGGCTCTACACCTTTCAATTCACATGGAATAACT 382
||||| |

Qy 797 actgaaagcttcgagggagctgaatatatgatacagaccagctcaagggagaaagaag 856
||||| |

Db 383 ACTTAAGAAGACATGAGGGGCCATTAACCTGTATACACCAACAGATTTTCAAGAGCAG 442
||||| |

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:24:21 ; Search time 5101.32 Seconds
(without alignments)
3237.138 Million cell updates/sec

Title: US-09-740-211-13_COPY_6000_7000

Perfect score: 1001
Sequence: 1 cctgcacttaccgcgcctc.....ggctgcttaactatgcgcg 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rnd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rnd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	2297	6 AX027819	AX027819 Sequence
2	1001	100.0	2368	12 AF264618	AF264618 Cloning v
3	1001	100.0	2417	12 CVPSP70	C65330 Cloning vec
4	1001	100.0	2419	12 CVPSP71	C65331 Cloning vec
5	1001	100.0	2450	12 CUV74374	U74374 Cloning vec
6	1001	100.0	2462	12 CVPSP72	C65332 Cloning vec
7	1001	100.0	2464	12 CVPSP73	C65333 Cloning vec
8	1001	100.0	2512	12 ASAU2679	AU002679 Synthetic
9	1001	100.0	2665	12 SYNPU88V	L08959 pUC8 clonin
10	1001	100.0	2665	12 SYNPU89V	L09128 pUC9 clonin
11	1001	100.0	2668	12 SYNPU87V	L08880 pUC7 clonin
12	1001	100.0	2674	12 SYNPU87V	L08958 pUC7 clonin
13	1001	100.0	2680	6 A02712	A02712 pUC13 DNA s
14	1001	100.0	2680	12 SYNPU812V	L09128 pUC12 clonin
15	1001	100.0	2680	12 SYNPU813V	L09130 Cloning vec
16	1001	100.0	2686	12 SYNPU818V	A02710 pUC18 DNA s
17	1001	100.0	2686	6 A02710	L08752 pUC18 clonin
18	1001	100.0	2686	12 SYNPU818V	M77789 Cloning vec
19	1001	100.0	2695	12 SYNPU819V	L08911 pUC119R clo
20	1001	100.0	2695	12 SYNPU819V	L08138 pUC8-1 expr
21	1001	100.0	2695	12 SYNPU822V	L09133 pUC8-2 expr
22	1001	100.0	2695	12 SYNPU822V	L09140 pUC9-1 expr
23	1001	100.0	2695	12 SYNPU822V	L09141 pUC9-2 expr
24	1001	100.0	2701	12 SYNPU819HV	L08910 pUC119H clo
25	1001	100.0	2704	12 SYNPU88V	L08941 pUC8 clonin
26	1001	100.0	2704	12 SYNPU89V	L08942 pUC8 clonin
27	1001	100.0	2710	12 SYNPU830V	L09142 pUC830 clon
28	1001	100.0	2710	12 SYNPU831V	L09143 pUC931 clon
29	1001	100.0	2713	12 SYNPU833V	L08939 pUC3 clonin
30	1001	100.0	2713	12 SYNPU84V	L08940 pUC4 clonin
31	1001	100.0	2716	12 SYNPU820RV	L08913 pUC20R clon
32	1001	100.0	2716	12 SYNPU821V	L08937 pUC1 clonin
33	1001	100.0	2722	12 SYNPU822V	L08938 pUC2 clonin
34	1001	100.0	2722	12 SYNPU820HV	L08912 pUC20H clon
35	1001	100.0	2743	12 CVPSEM4Z	X65304 Cloning vec
36	1001	100.0	2746	12 CVPSEM4Z	X65305 Cloning vec
37	1001	100.0	2854	12 CVPSEM4Z	U34922 Cloning vec
38	1001	100.0	2865	12 CVPSEM1	X65300 Cloning vec
39	1001	100.0	2865	12 SYNPU814V	L08870 pSEM1 clonin
40	1001	100.0	2867	12 CVPSEM3	X65302 Cloning vec
41	1001	100.0	2867	12 SYNPU833V	X65302 Cloning vec
42	1001	100.0	2869	12 CVPSEM2	X65301 Cloning vec
43	1001	100.0	2869	12 SYNPU822V	L08871 pSEM2 clonin
44	1001	100.0	2871	12 CVPSEM4	X65303 Cloning vec
45	1001	100.0	2891	12 CVP7NCO	U25273 Expression

ALIGNMENTS

RESULT 1
AX027819 2297 bp DNA PAT 16-SEP-2000
DEFINITION Sequence 11 from Patent WO0039307.
AX027819
VERSION AX027819.1 GI:10188663

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2297)
TITLE Bestetti, G., Cali, S., Orsini, G., Tonon, G., Zuffi, G. and Ghisotti, D.
Recombinant bacterial strains for the production of natural
nucleosides and modified analogues thereof

JOURNAL
BESTETTI GIUSEPPINA (IT) ; CALI SIMONA (IT) ; NORPARNA SPA (IT) ;
ORSINI GAETANO (IT) ; TONON GIANCARLO (IT) ; ZUFFI GABRIELE (IT) ;
GHISOTTI DANIELA (IT)

FEATURES
source location/Qualifiers
1..2297

/organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="cloning vector derived from pUC18"
 BASE COUNT 571 a 574 c 580 g 572 t
 ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 2297;
 Best local Similarity 100.0%; Pred. No. 1.2e-274;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacattatccgcctccatccagctcattatgttccggaagctagaagagt 60
 DB 1142 CCGCAACTTTATCCCGCTCCATCCAGCTATTAATTGTTCCGGGAGACGTAGAGTA 1083

QY 61 agttgcaggtatattgtttgcgaacgttgttgcattgtctacaggatctgtgtca 120
 DB 1082 AGTTGCCAGTTAATAGTTGGCCGAACGTTGGCCATTGTCAAGGCAATCGTGGTCA 1023

QY 121 cgtctcgttgtgtatgttgcattccagctccggtcccaacgatacaggagttaca 180
 DB 1022 CGCTCGCTGTTGGTATGGCTTCATTCACCTCCGTTCCCAAGATCAAGCGAGTTACA 963

QY 181 tgaatcccccattgttgcacaaaagcggttaagctccttcgtctccgcgtctgtcaga 240
 DB 962 TGATCCCGCATGTTGCAAAAAAGCGGTTAGCTCCTTCGCTCCGATCCGTTGTGACA 903

QY 241 agtaagttgcgagtgatcatcactcagtgatgagcagactgataatctctact 300
 DB 902 AGTAAGTTGGCCGAGTATACCACTAGTATAGGCGACACTGCATTAATCTCTTACT 843

QY 301 gtcattgcacccgtaagatgctttcttgactggtgagtaactcaaccaagctattcga 360
 DB 842 GTCATCCATCCGTAGATAGCTTTTCTGTGACTGGTGAATCAACCAAGCATTTCTTA 783

QY 361 gaatagtgatgcgcgcagagtgcttcttgccggcgctcaatacgggataatccgcg 420
 DB 782 GAATAGTGTATGGGGGACCGAGTTGCTCTCCCGGCTCATATCGGATTAATACCGCG 723

QY 421 ccacatagacagaacttaaaagtgcacatcatctggaagaacgtttcttgagggaagaactc 480
 DB 722 CCACATAGCAGAACTTTAAAAGTGCCTCATCTTGGAAAAGCTTCTTCGGGGGAAACCTC 663

QY 481 tcaagatcttaacgcgtgttgatccagtlcgatgtaaccactcgttgcacccaactga 540
 DB 662 TCAAGATCTTACCGCTTGTGAGATCCAGTTGATGTAACCACTCGTGACCACTGA 603

QY 541 tcttgcagctcttacttcaaccagcggttcttggttgagcaaaaacaggagcaaat 600
 DB 602 TCTTCAGCATCTTTTACTTTCACACGCGTTCTGGGTGAGCAAAAACAGGAAGCAAAAT 543

QY 601 gccgcaaaaagggaatlaaggcgacacggaatgttgaatactatctctcttctt 660
 DB 542 GCCGCAAAAAGGAATTAAGGCGACACGGAATTTGAAATCTATCTCTCTCTTTT 483

QY 661 caatatattgaagcattatcagggattatgttcacatgagcggaatacatattgaagt 720
 DB 482 CAATATTATTGAAGCATTTATCAGGGTTATGTCTCATGAGGAGTACATATTGAATGT 423

QY 721 atttagaaaaataaacaatagggttccgcgacattccccgaaaagtccacgtac 780
 DB 422 ATTTAGAAAAATAACAATAGGGGTTCCGGCACATTTCCCGAAAAGTCCACCTATAC 363

QY 781 gtctagaacaacctattatcatcagcattaaactlaaaataagcgtatatacaggcc 840
 DB 362 GTCTAGAAACCATATTATATCTGACATTAACCTATAAAATTAAGCGATACAGAGGCC 303

QY 841 ttctgctcgcgcttctcgttgaagaacgtctcagacatgcagctcccgag 900
 DB 302 TTTCTCTCGCGCTTTCGGGTGATGACGGTGAAACCTCTGACACATGCACCTCCGGAG 243

QY 901 acggtcacagcttctctgaagcgatgcgggagcagacaagccgctcaggcgcgctca 960
 DB 1099 AGTTCGCCAGTTAATAGTTGGCCGAACGTTGTCATTCAGGACATCGTGGTGTCA 1040

DB 242 ACGTCACAGCTTCTGTGAACGGATCCCGGAGACAGCAACCCGCTCAGCGCGTCA 183
 QY 961 gcgggtgttgcggggtgtcgcgggtcgtcgttactatgagc 1001
 DB 182 GCGGGTGTGGGGGTGTGCGGGCTGCTTAATCATATGCGGC 142

RESULT 2
 AF264618/c
 LOCUS AF264618 2368 bp DNA circular SYN 03-OCT-2000
 DEFINITION Cloning vector pUC6S, complete sequence.
 ACCESSION AF264618 M74308
 VERSION AF264618.1 GI:10504974
 KEYWORDS
 SOURCE
 ORGANISM
 Cloning vector pUC6S.
 Cloning vector pUC6S
 artificial sequence; vectors.
 REFERENCE
 1 (bases 1 to 2368)
 Vieira, J. and Messing, J.
 New pUC-derived cloning vectors with different selectable markers
 and DNA replication origins
 Gene 100, 189-194 (1991)

JOURNAL
 MEDLINE 91276241
 PUBMED 1905257
 REFERENCE
 2 (bases 1 to 2368)
 Vieira, J. and Messing, J.
 Direct Submission
 Submitted (27-APR-1993) Waksman Institute, Rutgers State
 University, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 3 (bases 1 to 2368)
 Vieira, J. and Messing, J.
 Direct Submission
 Submitted (26-SEP-2000) Waksman Institute, Rutgers State
 University, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Sequence update by submitter
 On Oct 3, 2000 this sequence version replaced gi:209216.
 LOCATION/Qualifiers
 1..2368
 1..2368

REMARK
 COMMENT
 FEATUERS
 source
 misc-feature
 1..127
 /db_xref="taxon:137702"
 1..127
 /note="multiple cloning site"
 513..1373
 /gene="bla"
 513..1373
 /gene="bla"
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="AAC18332.1"
 /db_xref="GI:10504975"
 /translation="MSIQHFRVALILFPAARCLPFAHPETLVKVKDAEDQLGARVGY
 IELDLSGKILFSFRPEERFPMWSDFKVLCAVLSRIDAGEQLGRIRHISQNDLVE
 YSPVTEKHLTDGTVARELCSAAITMSDNTAAMLLITIGPKRELTAFLHNMDSYRL
 DRWEPLNEAIFADDERDTTTPYVMAATTLRKLLTGLIELTLASROQLIDMWEADKVGPL
 LRSAIPAGWFIADKSGAGRGSRGRIIALGPDGKPSRIIVITVTSQATMDERNRQIA
 ELGASLIRKHM"

BASE COUNT 583 a 589 c 600 g 596 t
 ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2368;
 Best local Similarity 100.0%; Pred. No. 1.2e-274;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacattatccgcctccatccagctcattatgttccggaagctagaagagt 60
 DB 1159 CCGCAACTTTATCCCGCTCCATCCAGCTATTAATTGTTGCCGGGAGACTAGAGTA 1100

QY 61 agttgcaggtatattgtttgcgaacgttgttgcattgtctacaggatctgtgtgtca 120
 DB 1099 AGTTCGCCAGTTAATAGTTGGCCGAACGTTGTCATTCAGGACATCGTGGTGTCA 1040

```

QY 121 cgcctgcgttggatggtcattcgaagtcgggtcccaagcagatcaaggcgagtaca 180
DB 1039 CGCTGCTGTTGGATGCTTCAATCAGCTCCGCTTCCCAACATCAGGCGAGTTACA 980
QY 181 tgatcccccattggtgtaaaaaaagcggttagcccttcggtccctccatggtgtcaga 240
DB 979 TGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCCCTCCGATCGGTTGTCA 920
QY 241 agtaagttggcgcagtgltatcactcaatgtgtatgagcagactgcatatcttact 300
DB 919 AGTAAGTTGGCGCGAGCTGTATCACTCATGTTATGAGCAGCATGATTAATTTCTACT 860
QY 301 gtcacgcatccgtaagaatgtcttctgtgactgtgtgactcaaccaatcattctga 360
DB 859 GTCATGCCATCCGTAAAGATGCTTTCTGTGACTGCTGACTCAACCAAGTCATCTGA 800
QY 361 gaatagtatagcgcgcgcaggtgtccttcgcccgcgcataatcaggaataacgcg 420
DB 799 GAATAGTGTATGCGCGAGCAGCAGTGTCTTCCGCGGCTCAATACGGGATATACCGCG 740
QY 421 ccacatagcagacacttaaaagtgctcaatctggaanaagcttctcgggcgcaaac 480
DB 739 CCACATAGCAGAACTTTAAAGTGTCAATGGAAGAACTTTCTCGGGCGAAGAACTTC 680
QY 481 tcaagatcttcacgcgtgtgtgagatccagttcgatgtaaccactcgtgaccccaactga 540
DB 679 TCAAGATCTTACCGCTGTGAGATCCAGTGTGATGATACCCACTGTCGACCCACTGA 620
QY 541 tcttcagcacttcttacttccacgcgcttctgtgtgagcaaaaacaggaagcaaat 600
DB 619 TCTTCAGCATCTTTACTTTCACAGCGTTCTGTGGTGAGCAAAAACAGGAAGCAAAAT 560
QY 601 gccgcacaaaaggaataaggcgacacgcgaagaatgttgaataactactcttctt 660
DB 559 GCCGCAAAAAGGAATAAGCGGACGCGAAGTGTGATTAATCTACTCTCTCTTTT 500
QY 661 catatattgaagcattatcatcaggttatgttccatgaagcggtatcatatttgaat 720
DB 499 CATATTATTGAAGCTTTATCAAGGTTATTTCTCATGAGCGGATACATATTAAATGT 440
QY 721 atttaaaaaataaacaataagggttccgcgcacattcccgcaaaagtgcaccctgac 780
DB 439 ATTAAACAAATAAACAATAAGGGGTTCCGCGCATTTCCCGCAAAAGTGCACCTGAC 380
QY 781 gtctaaagaacattatcatcagcattacactataaagaagcggtatccgcgagcc 840
DB 379 GTCTAAGAAACCATTAATATGATGACATTAACTTAATAAATAGCGGTATCAGAGGCC 320
QY 841 ttctcgtcgcgcgttgggtgatagcggtgaanaacctgtgacatgagctccgcgag 900
DB 319 TTTCTGCTCGCGGCTTGTGATGATGAGGTTGAAAACCTGTGACATGACGCTCCCGAG 260
QY 901 acggtacagcctgtctgtaagcgaatgcgggagcagaacacccgctcagggcggtca 960
DB 259 ACGGTACAGCCTTGTGTAAAGGAGATGCCGGGAGCAGACAGCCCGTACGGCGCTCA 200
QY 961 ggcgggtgttgccggggtgtcgaggctgagcttaactatgagc 1001
DB 199 GCGGCTGTGTGCGGCTGTGCGGCTGCTTAACATAGCGGC 159

```

```

RESULT 3
CVPSP70 2417 bp DNA circular SYN 16-JUL-1999
LOCUS Cloning vector psp70.
DEFINITION X65330
ACCESSION X65330.1 GI:58235
VERSION beta-lactamase; bla gene; cloning vector; multiple cloning site;
KEYWORDS Promoter.
SOURCE Cloning vector psp70.
ORGANISM Cloning vector psp70.
REFERENCE 1 (bases 1 to 2417)

```

```

AUTHORS Technical Services.
TITLE Direct Submission.
JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
REMARK Revised by 21
REFERENCE 2 (bases 1 to 2417)
AUTHORS Technical Services.
TITLE Direct Submission.
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
See X65300-X65335 for related vector sequences
This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330.
FEATURES
source
location/Qualifiers
1..2417
/organism="Cloning vector psp70"
/db_xref="taxon:90135"
misc_feature
1
/note="Sp6 transcription initiation site"
misc_feature
4..45
/note="multiple cloning sites"
misc_feature
56
/note="T7 transcription initiation site"
57..73
/note="T7 promoter"
gene
complement(1090..1950)
/organism="bla"
complement(1090..1950)
/organism="bla"
/codon_start=1
/transl_table=11
/product="Beta-lactamase"
/protein_id="CAA46430.1"
/db_xref="GI:58236"
/transl_table="MSI0HFRVALIPFPAFLCPVFAHETLVKXKADBDLGARVY
IELDNSGKILSFREPERPMSTKVLICGAVLSRIADGOLFGRIRHYSQNDLYE
KSPVTEKHITDQMTVRELSAATIMSDNIAANLLTTIGPKELFAFLHMGDHTRL
DRNEPELNEAIIPNDRDTTPVPAAMATTKRLITGELLTLASROILDMWEADKVAAPL
LRASLPAGWEIANKSAGSGEKSGITIAALPDSKPSRIYIYITTSQATIMDERNRRIA
EIGASLTKHM"
promoter
2401..2417
/note="Sp6 promoter"
BASE COUNT 615 a 599 c 592 g 611 t
ORIGIN
Query Match 100.0%; Score 1001; DB 12; Length 2417;
Best Local Similarity 100.0%; Pred. No. 1,2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cctgaacttataccgctccatccagctcattatgttgcgggaagctagagtaagt 60
DB 1304 CCGCAACTTATCCGCTCCATCCAGCTATTAATGTTCCCGGAAGCTAGAGTAAGT 1363
QY 61 agtcgcagtaagaatttgcgaagctgttgcacattgcttaccagagatgtgtgtca 120
DB 1364 AGTTCGCCAGTTAATAGTTTGCGCAACGTTGTGCAATGCTACAGGCAATCTGTGTGCA 1423
QY 121 cgcctgcgttggatggtcattcgaagtcgggtcccaagcagatcaaggcgagtaca 180
DB 1424 CGCTGCTGTTGGATGCTTCAATCAGCTCCGCTTCCCAACATCAGGCGAGTTACA 1483
QY 181 tgatcccccattggtgtaaaaaaagcggttagcccttcggtccctccatggtgtcaga 240
DB 1484 TGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTCCCTCCGATCGGTTGTCA 1543
QY 241 agtaagttggcgcagtgltatcactcaatgtgtatgagcagactgcatatcttact 300
DB 1544 AGTAAGTTGGCGCGAGCTGTATCACTCATGTTATGAGCAGCATGATTAATTTCTACT 1603

```



```

QY 301 gtcacgcatcgttaagatgcttctgtgactggtgagtaactcaacgaagtcattcga 360
    |||||||
Db 1604 GTCATGCCATCCGTAAGATCTTTTCTGTGACTGGTAGTCAACCAAGTATTCTGA 1663
QY 361 gaatagtgtatgcygcagccaggtgtctctgtccgcgcgtcaatacaggaataacagcg 420
    |||||||
Db 1664 GAATAGTGTATGCGGCGACGAGTGTCTTGGCCGCGCTCAATACGGATATACCGGG 1723
QY 421 ccacatagcagaacttaaaagtgctcattatgtgaaaagcttctcggggcgaaactc 480
    |||||||
Db 1724 CCACATAGCAGAACTTTAAAGTGTATCATTTGAAACGTTCTTGGGCGCAAAACATC 1783
QY 481 tcaaggatctaccgctgttgagatccagttcagtgatgaaccactggtcacccaactga 540
    |||||||
Db 1784 TCAAGATCTTACCGCTGTGAGATCCAGTTCATTAACCACTGTCACCCAACTGA 1843
QY 541 tcttcagcatcttcttacttaccagcgcttctcgggtgagcaaaaacaggaagcaaat 600
    |||||||
Db 1844 TCTTACGACATCTTTTACTTTCACACGCTTTCGGGTGAGCAAAACAGGAAGCAAAAT 1903
QY 601 gccgcaaaaaagggaataaggcgagacaggaatgttgatactactcttccctttt 660
    |||||||
Db 1904 GCCGCAAAAAAGGAAATTAAGGCGACAGGAATGTGAATCTCATCTCTTCCCTTTT 1963
QY 661 caatattatgaagcatctatcaggggtatgttctcatgagcgatatactattgaatgt 720
    |||||||
Db 1964 CAATATTATTGAAGCATTTATCAGGCTTATGTCATGAGCGGATATCATATTGAAATGT 2023
QY 721 attagaataaataaagaataagggtgtccgcagacattccccgaaaagtgccactgac 780
    |||||||
Db 2024 ATTTAATAAATAAACAATAAGGGGTTCGCGCACATTTTCCCGAAAAGTGCCACCTGAC 2083
QY 781 gctcaagaacacattatcatcagacataacctaataaagaagcgatcagagggccc 840
    |||||||
Db 2084 GCTAAGAACAACATTATTCATGACATTAACCTATAAATAAGGCGTATCAGAGGCC 2143
QY 841 ttctgtctgcgcgttctggtgtatgaggtgaaaacctctgacacatgcagctcccgag 900
    |||||||
Db 2144 TTTGCTGTGCGCGTTCGCGTATGACGTGAACACCTCTGACATGACAGCTCCCGAG 2203
QY 901 acggtacagctgtctgtgaagcagatgccggagacagaacagccctcagagcgcgca 960
    |||||||
Db 2204 ACGGTACAGCTTGTCTGTGAACGAGTCCGGGACAGCAAGCCCGTCAAGGCGCGTCA 2263
QY 961 gcggggtgtgtgcgggtgtctgcgggctgacttaactatgcggc 1001
    |||||||
Db 2264 GCGGGTGTGCGGGGTGTCGGGGCTGACTTAATGCGGCG 2304

RESULT 4
LOCUS CVPSP71 2419 bp DNA circular SYN 16-JUL-1999
DEFINITION Cloning vector pSP71.
ACCESSION X65331
VERSION X65331.1 GI:58237
KEYWORDS beta-lactamase; bla gene; cloning vector; multiple cloning site;
SOURCE
ORGANISM Cloning vector pSP71.
REFERENCE Cloning vector pSP71.
AUTHORS artificial sequence: vectors.
TITLE 1 (bases 1 to 2419)
JOURNAL Technical Services.
REMARK Direct Submission.
AUTHORS Submitted (23-MAR-1992) Technical Services, Promega Corporation,
TITLE 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
JOURNAL revised by (2)
REMARK 2 (bases 1 to 2419)
AUTHORS Technical Services.
TITLE Direct Submission.
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,
COMMENT 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
See X65300-X65335 for related vector sequences

```

This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical information:

Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330.

FEATURES

source

1..2419 location/Qualifiers

/organism="Cloning vector pSP71"

/db_xref="taxon:90136"

misc_feature

1 /note="Sp6 transcription initiation site"

misc_feature

6..47 /note="multiple cloning sites"

promoter

53..75 /note="T7 promoter"

misc_feature

58 /note="T7 transcription initiation site"

gene

complement(1092..1952)

CDS

/gene="bla"

complement(1092..1952)

/gene="bla"

/codon_start=1

/transl_table=11

/product="Beta-lactamase"

/protein_id="CAA46431.1"

/db_xref="GI:58238"

/translation="MSIQHFRVALIIFPAFLCPLPAFHEPLLYKKVDAEDQSGARVY

IEIDNSGKLTLESFREERPMSTKRVLCGVLSRIDAGEOLGRTHYSQNDLVE

YSPVTEKHLLIDGTVRELCSALTMSDNTAANLLTITGPELFAFLNMDDHYRL

DRWEPELNEAIPNDERDTPMPVAMATYLRKLITGELTLASROQLIDWMEADKVGAPL

LRSLAPGMEITADKSAGERSRGIIAALGPDPKPSRIYVITTSQATMDERNQIA

EIGASLIKHW"

2402..2419

promoter

/note="Sp6 promoter"

BASE COUNT

615 a 601 c 592 g 611 t

ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2419;

Best Local Similarity 100.0%; Pred. No. 1,2e-274;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgcaacttlatccgctccatccatccatcattatgtgtgcgggaagcagaagtaagt 60
    |||||||
Db 1306 CCTGCACTTTATCCGCTCCATCCAGTATTAATGTTGCCGGGAGAGCTAAGTAAGT 1365
QY 61 agttcgccaglttaagatttgcgaacggtgtgtccattgtctacaggaagcatcgtgtgca 120
    |||||||
Db 1366 AGTTCGCCAGTTAAATGTTGCGCAACGTTGTCATTCCTACAGGCATCCGTGTGTCA 1425
QY 121 cgtcgtcgttggtaagctcattcagctcaggttcccaagatccaagcgagttaca 180
    |||||||
Db 1426 CGCGTCGTCTTGGTAGAGCTTATTCATTCAGCTCCGGTCCCAACGATCAAGCGAGTTTACA 1485
QY 181 tgaatcccaatgttgcgaataaagcggtagctccttgcgtccttcgaatcgtgtgcaga 240
    |||||||
Db 1486 TGATCCCCCATGTTGTGCAAAAGGCGTTAGCTCTCGTGGTCCGATCGTGTGCAGA 1545
QY 241 agtaagttgacgagtgatgtatatacctcaatggtatagcgagcatgtatatactctact 300
    |||||||
Db 1546 AGTAAAGTTGGCCGACAGTGTATCATCATGTGTTATGACATGATGATTAATCTCTTACT 1605
QY 301 gtcacgcatcgttaagatgcttctgtgactggtgagtaactcaacgaagtcattcga 360
    |||||||
Db 1606 GTCATGCCATCCGTAAGATCTTTTCTGTGACTGGTAGTCAACCAAGTATTCTGA 1665
QY 361 gaatagtgtatgcygcagccaggtgtctctgtccgcgcgtcaatacaggaataacagcg 420
    |||||||
Db 1666 GAATAGTGTATGCGGCGACGAGTGTCTTGGCCGCGCTCAATACGGATATACCGGG 1725
QY 421 ccacatagcagaacttaaaagtgctcattatgtgaaaagcttctcggggcgaaactc 480
    |||||||

```

Db 1726 CCACATAGCAGCACTTTAAAGTGCATCATTTGGAAGAAAGCTTCTTGGGGGCGAAAACTC 1785
 QY 481 tcaagagatctaccgctgttgagatccagttcgatgtaacccaactcgtgacccaactga 540
 Db 1786 TCAAGGATCTTACCGCGTGTGAGATCCAGTTCAGTGAACCCACTCGTGCAACCACTGA 1845
 QY 541 tcttcacacatcttactcttcacacagcgtttctggtgagcaaaaaggaagcaaat 600
 Db 1846 TCTTCACGATCTTTACTTTCACACGCGTTCTGCGTGAGACAAACAGAGAGCAAAAT 1905
 QY 601 gccgcaaaaaggaagcaagggcgacacaggaatggtgaatactactactcttctttt 660
 Db 1906 GCCGCAAAAAGGGAATTAAGGCGACACAGAAATGTTGAATACATCATCTTCTCTTTT 1965
 QY 661 caataattgaagcattatcaggggtatgtctctatgagcgagatacatattgaatgt 720
 Db 1966 CAATATTATTGAACATTTATCAGGGTTATTTGTCTCATGAGCGGATCAATATTGAAATGT 2025
 QY 721 attgagaaaaaataacaataaggggttcgcgcgcacattcccccgaagaagtcgac 780
 Db 2026 ATTGAGAAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTCCACCTGAC 2085
 QY 781 gtcaagaacacattatcatatgacatacctaataaagaagcgatcacgaagccc 840
 Db 2086 GTCTAAGAAACCATTTATATCATGACATTAACCTATAAAATAGCGGTATCAGAGGCC 2145
 QY 841 ttctgcctcgcgcttcggtgagtgatgagtgagaaacgttcgacatgacgctccgag 900
 Db 2146 TTTCGTCGCGCGCTTTCGTTGATGACGAGTGAAGAAACCTGACACATGAGCTCCCGAG 2205
 QY 901 acgttcacagcttctgttaagcgagatgcgcggagcgagacagacccgtcagggcgctca 960
 Db 2206 ACGGTCACAGCTTGTGTGTTAGCGGATGCGGGAGCAGACAGCCCGTCAAGGCGCTCA 2265
 QY 961 gcgggtgtgagcggtgtgcggaggtgcgttaactatgagc 1001
 Db 2266 GCGGCTGTGGCGGGTGTGCGGGCTGGCTTAATGATGCGGC 2306

RESULT 5
 CVU74374 2450 bp DNA circular SYN 29-OCT-1996
 LOCUS Cloning vector pJF5, complete sequence.
 DEFINITION U74374
 ACCESSION U74374
 VERSION U74374.1 GI:1649038
 KEYWORDS Cloning Vector pJF5.
 SOURCE Cloning Vector pJF5.
 ORGANISM Cloning Vector pJF5.
 REFERENCE 1 (bases 1 to 2450)
 AUTHORS Fan, J.-B., Quackenbush, J. and Myers, R.M.
 TITLE pJF5 Cloning Vector
 JOURNAL Unpublished (1996)
 REFERENCE 2 (bases 1 to 2450)
 AUTHORS Fan, J.-B., Quackenbush, J. and Myers, R.M.
 TITLE Direct Submission
 JOURNAL Submitted (11-OCT-1996) Stanford University, Stanford Human Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA
 COMMENT Contact: R. M. Myers
 Stanford Human Genome Center
 855 California Ave.
 Palo Alto, CA 94304 USA.
 Ampicillin Resistance Selectable Vector used in Transposon-Mediated Sequencing.
 FEATURES
 source
 1. 2450
 /location/Qualifiers
 /organism="Cloning Vector pJF5"
 /db_xref="taxon:53561"
 misc_feature
 1
 /note="Sp6 transcription initiation site"
 misc_feature
 4. 78
 /note="Multiple Cloning Site
 XhoI/PvuII/BstXI/XbaI/BamHI/EcoRI/BstXI/EcoRV/BglII"

promoter 86..106
 /note="T7 promoter"
 misc_feature 88
 /note="T7 transcription initiation site"
 CDS complement(1123..1983)
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="AAB17670.1"
 /db_xref="GI:1649039"
 /translation="MSIQHFRVALIPFAKCLIPFAHPHTLVKQAEQDGLARVGY
 IELDINSKILSEFRPEERFPMSTFKVLICGAVLSRIDAOBOLGRIRIHQSNDLVE
 YSPVTEKHLDGMVRELCSAIIIMSDVTANLLITTTGKRELTAFLHNGDVTRL
 DRMEDELNDIIPNDERDTMPVAMATYTRKILITELTLASRODLIDMEADKVAQPL
 LRSALPAGMTADKSGAGERSRGIIALGPDGKRSRIVIYITGSOATMERNRQIA
 ETGASLIRKH"
 promoter 2434..2450
 /note="Sp6 promoter"
 BASE COUNT 624 a 609 c 602 g 615 t
 ORIGIN
 Query Match 100.0%; Score 1001; DB 12; Length 2450;
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacattatccgcctccatccagttctattatgttgcgggaagctagagtaagt 60
 Db 1337 CCGCAACTTATACCCCTCCATCCAGTCAATTAATGTTGCGGAGACTAGATTAAGT 1396
 QY 61 agtgcacagttatagttgtgcgaacgtgtgtccattgtctacaggaacgtgtgtgtca 120
 Db 1397 AGTTCGCATTAATAGTTTGGCGCAAGTTGTTGCCATTGCTACAGGACGCGGTGCA 1456
 QY 121 cgtctcgctgtgtgtatggtctatccagtcgcggtcccaagatcaagcgagttaca 180
 Db 1457 CGCTCGCTGTTGTGTAATGGCTTCATTCAGCTCCGCTCCCAAGATCAAGCGAGTTACA 1516
 QY 181 tgaatcccatgtgtgtgcgaacaaagcggttaagtccttcgttcctcgatcgtgtgcaga 240
 Db 1517 TGAATCCCATGTTGTGTCAAAAAGCGGTAGCTCTCTCGGTCTCCGATCGTTGTCA 1576
 QY 241 agtaagttgcgcagttgtatcatcaatagttatgagcagcactgataatctcttact 300
 Db 1577 AGTAAGTTGGCGCAGGTATCATCATCATGTTATGACACACTGCATTAATCTCTTACG 1636
 QY 301 gtaatgcacatcgtaagaagctttctgtgacgtgtgtgaggtacccaagatattctga 360
 Db 1637 GTCATGCCATCCCTAAGATCCTTTCTGTGACGTGTGAGTACCAACCAAGTCTATTCTGA 1696
 QY 361 gaatagtatagtgagcgacaggttgccttgccgcggcgatcaataggaataatccgcg 420
 Db 1697 GAATAGTGTATGCGGCGACCGAGTTGCTCTTGGCCGCGCATACGGGATTAATACGGG 1756
 QY 421 ccaatagcagaacttaaaagtgctcatcatatggaanaagcttctcggggcgaaactc 480
 Db 1757 CCACATAGCAGAACTTAAAGTGCATCATGTAAGAAAGTCTTCCGGGCGAAAACTC 1815
 QY 481 tcaagagatctaccgctgttgaaatccagttcgatgtaacccaactcgttgacccaactga 540
 Db 1817 TCAAGGATCTTACCGCGTGTGAGATCCAGTTCAGTGAACCCACTGTGTGCCAACACTGA 1876
 QY 541 tcttcacacatcttacttaccacagcgtttctggttgaagcaaaaaggaagcaaat 600
 Db 1877 TCTTCACGATCTTTACTTTCACACGCGTTCTGCGTGAGCAAAAACAGAGAGCAAAAT 1936
 QY 601 gccgcaaaaaggaagcaagggcgacacaggaatggtgaatactactactcttctttt 660
 Db 1937 GCCGCAAAAAGGGAATTAAGGCGACACAGAAATGTTGAATATCATCTTCTCTTTT 1996
 QY 661 caataattgaagcattatcaggggtatgtctctatgagcgagatacatattgaatgt 720
 Db 1997 CAATATTATTGAACATTTATCAGGGTTATTTGTCTCATGAGCGGATCAATATTGAAATGT 2056

QY 721 attagaaaaataacaataggggttcgcgcacatttcccgaaaaagtgcacctgac 780
|||||
Db 2057 ATTTAGAAAATTAACAAATAGGGGTTCCCGCACAATTTCCCGAAAGTGCCACCTGAC 2116
QY 781 gctaaagaacattatcatcatgacattacattataaataagcgctatcagagccc 840
|||||
Db 2117 GTCTAAGAAACATTATTATCATGACATTAACCTATTAATAATAGGGGTATCAGAGGCC 2176
QY 841 ttctgctgcgcgtttcgcgtatgacggtgaaacctctgacacatgcagctccggag 900
|||||
Db 2177 TTTGCTTCGCGGCTTGGTGGATGACGGGTGAACCTCTGACACATGCACTCCCGGAG 2236
QY 901 acggtcacactgtctgttaagcagatgcgcgagcagacacccgtcagggcgagta 960
|||||
Db 2237 ACGGTACACACTTCTCTGTAAAGCGATGCGGAGACGACACACCCGCTCAGGGCGCTCA 2296
QY 961 gcgggtgtgcgcgggtgcgcgggtgctgcttaactatgcgc 1001
|||||
Db 2297 GCGGGTGTGGCGGCTGCGGGCTGCTTAACCTATGCGGC 2337

RESULT 6
LOCUS CVPSP72 2462 bp DNA circular SYN 25-JAN-2000
DEFINITION Cloning vector pSP72.
ACCESSION X65332
VERSION X65332.2 GI:6759494
KEYWORDS beta-lactamase; bla gene; cloning vector; multiple cloning site;
promoter.

SOURCE Cloning vector pSP72.
ORGANISM Cloning vector pSP72.
REFERENCE 1 (bases 1 to 2462)
ARTIFICAL SEQUENCE: vectors.
AUTHORS Technical Services.
TITLE Direct Submission.
JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA

REMARK 2 (bases 1 to 2462)
AUTHORS Technical Services.
TITLE Direct Submission.
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA

REMARK 3 (bases 1 to 2462)
AUTHORS Technical Services.
TITLE Direct Submission.
JOURNAL Submitted (12-JAN-2000) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA

COMMENT On Jan 26, 2000 this sequence version replaced gi:58239.
See X65300-X65335 for related vector sequences
This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330.

FEATURES
source
1. 2462
/organism="Cloning vector pSP72"
/db_xref="taxon:90137"
misc_feature
1
/note="SP6 transcription initiation site"
misc_feature
4..90
/note="multiple cloning sites"
promoter
99..118
/note="T7 promoter"
misc_feature
101
/note="T7 transcription initiation site"
gene
complement(1135..1995)
/gene="bla"
CDS
complement(1135..1995)

/gene="bla"
/codon_start=1
/transl_table=11
/product="Beta-lactamase"
/protein_id="CAA6432.1"
/db_xref="GI:58240"
translation="MSIOHFRVALIPFEAFCLPVPFAHPEFTLVKRADEPOLGARYX
IEDLNSKILIESFRPEPRPMSTFVKVLGAVLSRIDGQDOLGRIRHYSNDLYE
YSPYTERKHLTDGRTVRELCSAITSMSNTANLLTTIGPKELTAFIDMKGIVRL
DRWEPELNEALPNDERTTTPVAAATLRLKLTGELLTLASROOLDMWEADVVAGPL
LRSAIPAGWFIADKSGAGERSRGIILALPDGKPSRIIVYITTSQATDERNRQIA
EIGASLIKHW"
join(2446..2462,1..3)
promoter
/note="SP6 promoter"
BASE COUNT 624 a 614 c 605 g 619 t
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2462;
Best local similarity 100.0%; Pred. No. 1,2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaacttattccgcctccatccagctatattgttgcgggaagctagaagta 60
|||||
Db 1349 CCGTCACTTATCCGCCCTCCATCCAGTCTATTAATTGTCCGGGAGTAGAGTAGT 1408
QY 61 agttgcagtaattagtttgcgaagctgttgcattgtctaaagcattgtgttca 120
|||||
Db 1409 AGTTGCCAGTTAATAGTTTGGCAACGTTGTGCTTGCATGCTACAGCATCTGTGTCA 1468
QY 121 cgtctcgtttgttattgcttcatcagctcgcgttcccaacagatcaagcgagttaca 180
|||||
Db 1469 CGCTGTGTTGGTATGAGTTCATATGACGTCCGGTTCCCAACATCAGGCGAGTTACA 1528
QY 181 tgatccccaatgttgcgaagcggttagctcccttgcgtccatccgattgttgcaga 240
|||||
Db 1529 TGATCCCAATGTTGTGCAAAAAGCGGTTAGCTTCCTCCGATCGTTGTGAGA 1588
QY 241 agtaagttgcgcgaagtttcatcctatgtttagcagacgcatcttccctact 300
|||||
Db 1589 AGTAAATTGGCCGAGTGTATTCATCTATGCGATGAGCATGATTAATCTCTTACT 1648
QY 301 gtcatgccatccgtaagaatgtcttcttgtagctgttgaactcaaccagatccttca 360
|||||
Db 1649 GTCATGCCATCCGTAAGATGCTTTGTGTGACTGTAGTCTCAACCAAGTCACTTCTGA 1708
QY 361 gaatagttattgcgcgcgaagttgtcttgcgcgcgctcaatacgggataatccgcg 420
|||||
Db 1709 GAATAGTGTATGCGCGCACGAGTTGCTTGGCCCGCTCAATACGGGATTAATACCGCG 1768
QY 421 ccaatagcagaactttaagaatgtctcatcatatggaaaaagttcttgcggcgaaaaatc 480
|||||
Db 1769 CCACATACCAAGACTTTAAAGTCTCATCATCTTGGAAGACGTTCTTGGGGCGCAAAATC 1828
QY 481 tcaagatccttacgcgtttagatccagttcgaatgtaaccactctgtcacccaactga 540
|||||
Db 1829 TCAAGGATCTTACCGCTGTAGATCCAGTTGATGTAACCCACTCTGCAACCACTGA 1888
QY 541 tcttcagcattcttacttcaaccagcttcttgcgttgcgaacaaaggaagcaaat 600
|||||
Db 1889 TCTTCAGCATCTTACTTTCACCCAGCGTTTCTGGGAGCAAAACAGGAAGGCAAAAT 1948
QY 601 gccgcaaaaaagggaataaggcgacacggaataatgtgaatactcaactcttcctttt 660
|||||
Db 1949 GCCGCAAAAAGGGAATPAAGGCGACAGGAATGTTGAATCTCATACCTCTCTTTT 2008
QY 661 caatattgaagcatttcaaggttatgtctcctcagcgagatatacatattgaatgt 720
|||||
Db 2009 CAATATTATTCAGCATTTTATCAGGTTATTGTCTCATGAGCGGATACATATTGTAATGT 2068
QY 721 attagaaaaataacaataggggttcgcgcacatttcccgaaaaagtgcacctgac 780
|||||
Db 2069 ATTTAGAAAATTAACAAATAGGGGTTCCCGCACAATTTCCCGAAAGTGCCACCTGAC 2128

QY 781 gctcaagaacattatatacatatcaacataaataagcgatcagagccc 840
 Db 2129 gctcgaagAACCATATATATCATACATTAACCTATAAAATAGCGCTATACAGAGGCC 2188
 QY 841 ttgcctccgcgcgttcctgagatgacgagtgaaacccctcaccacagcagctcccgag 900
 Db 2189 ttgcctccgcgcgttcctgagatgacgagtgaaacccctcaccacagcagctcccgag 2248
 QY 901 aacgtcaacgttctcgttaagcagatccgcgagcagacagaccgctcagggcgctca 960
 Db 2249 ACGGTCAACAGCTTGTCTGTATAGCGGATCCGCGAGACAGACAGCCCGTCAAGGCGCGTCA 2308
 QY 961 gcggagtgctgagcggagtgctcgagcgtgcttaactatgcgc 1001
 Db 2309 GCGCGTGTGGCGGCTGTGCGGCGCTTAACCTATGCGCG 2349

RESULT 7
 CVPSP73 2464 bp DNA circular SYN 25-JAN-2000
 LOCUS Cloning vector pSP73.
 DEFINITION X65333
 ACCESSION X65333.2 GI:6759495
 VERSION beta-lactamase; bla gene; cloning vector; multiple cloning site;
 KEYWORDS promoter.
 SOURCE Cloning vector pSP73.
 ORGANISM Cloning vector pSP73
 REFERENCE 1 (bases 1 to 2464)
 TITLE Technical Services.
 JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 REMARK 2 (bases 1 to 2464)
 REFERENCE Technical Services.
 JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 REMARK 3 (bases 1 to 2464)
 REFERENCE Technical Services.
 JOURNAL Submitted (12-JAN-2000) Technical Services, Promega Corporation,
 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
 COMMENT On Jan 26, 2000 this sequence version replaced
 See X65300-X65335 for related vector sequences
 This vector can be obtained from Promega Corporation, Madison, WI.
 Call one of the following numbers for order or technical
 information:
 Order or Technical 800-356-9526
 In Wisconsin 800-356-9526
 Outside U.S. 608-274-4330.

FEATURES
 source
 1. .2464
 /organism="Cloning vector pSP73"
 /db_xref="taxon:90138"
 misc_feature
 1
 /note="SP6 transcription initiation site"
 misc_feature
 6. .92
 /note="multiple cloning sites"
 promoter
 101. .120
 /note="T7 promoter"
 103
 /note="T7 transcription initiation site"
 gene
 complement(1137. .1997)
 /gene="bla"
 complement(1137. .1997)
 /gene="bla"
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 CDS

/protein_id="CAA46433.1"
 /db_xref="GI:58242"
 /translation="MSIOHFRVALLIPFAAFCLPVEAFHETLVKKAEDOLGARVGY
 IELDINSKATLESFRPEREPRMSPFKVILGCAVLSRIDAGEOLGRRIHYSONDLVE
 YSPVTEKHLTDMETVRELCSAATITMSDNTANLLFTTIGCKRELFAPLHNGDHYRL
 DWAEPELNAIPENDERDITPMVAWTTLRKLTEGLTLTLASRQOLIDMEADKAGPL
 LRSALPAGWFIADKSGAGERSGRIILAIGDPKPSRIIVYITTSQATMDERNRQIA
 EIGASLRKHW"
 join(2448. .2464,1..3)
 /note="SP6 promoter"
 BASE COUNT 624 a 615 c 606 g 619 t
 ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2464;
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttaccgcgcctccatccagtcctatcatgtgtgcccgggaagctagaagtg 60
 Db 1351 CCTGCAACTTATTCGCCGCTCCATCCAGTCTATTAATGTTGCCGGAGAGCTAGATAAGT 1410
 QY 61 agtcgcagtaataagttgagcaacgttggtgcaatgctacagagcatgtgtgtca 120
 Db 1411 AGTTCGCCAGTATTAAGTTGCGCAACGTTGTTGCCATTCCTACAGGCAATCGGTGTCA 1470
 QY 121 cgtctcgtcgtttgtagtggttcattcagctcgggttcccaagatcaagggagttaca 180
 Db 1471 CGCTCGTGTGTGTATAGGCTCAATTCAGCTCCAGGTTCCACAGATCAAGCGAGTTACA 1530
 QY 181 tgaatcccatgtgtgcaaaaaagcggtagtctcctcgtgctcccgatcgtgtgtca 240
 Db 1531 TGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGGCTCCGATCGTGTGACGA 1590
 QY 241 agtaagttgcccagcagtggtatacatcattgttatgagcagcatgataattcttact 300
 Db 1591 AGTAAGTTGGCCGCGAGGTATACCTCATGTGTTAGGCAACCTGCATTAATTCCTACT 1650
 QY 301 gtcatgcatccgtaagatcttctcgtgacagtgtagtactcaaccaagtcattcga 360
 Db 1651 GTCATGCCATCCGTAACATGCTTTCTGTGACAGTGTAGTACCAACCAAGTCATCTTGA 1710
 QY 361 gaatagtgatgcccgcagcagagtgctctgcccgcgctcaataagggataaccgcg 420
 Db 1711 GAATAGTGATAGCGCGCACCGAGTGTGCTTGCCTGCGGCGTCAATACCGGATACCGCG 1770
 QY 421 ccacatgacgaactttaaagtgctatcatcattggaataagttcttggggcgaaaaatc 480
 Db 1771 CCACATAGCAACACTTAAAGTGCTCATCATGGAAGAAAGTCTTGGGGCGAAAACTTC 1830
 QY 481 tcaagatcttaccgctgtgagatccagttcagatgtaaccacactgtgacccaactga 540
 Db 1831 TCAAGATCTTACCGCTGTGTAGATCCAGTTCGATGTAAACCACTCGTCACCCAACTGA 1890
 QY 541 tcttcagatcttacttcaacttcaaccagcttctggtgtgagcaaaaaggaagcaaat 600
 Db 1891 TCTTCAGATCTTACTTTCACCCAGCGTTCTGTGGGAGCAAAAAAGCAAGCAAAAT 1950
 QY 601 gccgcaaaaaagggaataaggcgacacggaatggtgaatacctacaccttcttctt 660
 Db 1951 GCCGCAAAAAAGGAATAAGGCGACACGGAAATGTTGATATCAATCAATCTTCTTTT 2010
 QY 661 caatatatgaagcatltaacaggttatgtgtcctcagtgagcggatataattgaatgt 720
 Db 2011 CAATATTTTGAAGCATTTATCAGGGTATGTCTCATGAGCGGATACATATTTGAATGT 2070
 QY 721 attagaataaatacaaataggggttcgcgcacattcccgaaaaagcgacacgtgac 780
 Db 2071 ATTAGAATAAATAAACAATAGGGGTTCCGCGACATTTCCCAAAAAGTCCACCTGAC 2130
 QY 781 gctcaagaacattatatacatatcaacataaataagcgatcagagccc 840
 Db 2131 GCTCAAGAACCATTATATCATACATTAACCTATAAAATAGGCGTATCAGAGGCC 2190

QY 841 ttctgtctcgcgcgtttcgttgatgaagcgtgaacacctcgtacacatgacgtcccgag 900
Db 2191 TTTCGTTCGGCGGTTTCGGTGTGATGACGGTGAAGAACTCTGACATGCACTCCGGGAG 2250
QY 901 acggtacacagcttctgttaagcgaatgcgcggagacagacacccgtaaggcgatca 960
Db 2251 ACGGTACACAGCTTGTCTGTAAAGCGATGCGCGAGACAGAACCCGTCAGGCGCGTCA 2310
QY 961 gcgggtgttgccgggtgttcggcggtgactgaactatgagcg 1001
Db 2311 GCGGGTGTGCGGGGTGTGCGGGGTGCTTAAGTATGCGGC 2351

RESULT 8
ASAJ2679 2512 bp DNA SYN 28-MAR-2000
LOCUS ASAJ2679
DEFINITION Synthetic DNA containing Escherichia coli plasmid pFaba, bla gene,
multiple cloning site, pSP72 derivative.
ACCESSION AJ002679
VERSION AJ002679.1 GI:2623973
KEYWORDS beta-lactamase: bla gene.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2512)
AUTHORS Wach, A., Brachet, A., Pohlmann, R. and Philippsen, P.
TITLE New heterologous modules for classical or PCR-based gene
disruptions in Saccharomyces cerevisiae
JOURNAL Yeast 10 (13), 1493-1808 (1994)
MEDLINE 9526357
REFERENCE 2 (bases 1 to 2512)
AUTHORS Philippsen, P.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1997) Philippsen P., Applied Microbiology,
University of Basel, Biozentrum, Klingelbergstr. 70, Basel,
CH-4056, SWITZERLAND

FEATURES
source
1. 2512
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature
5. 134
/note="Multiple cloning site"
gene
complement(1185..2045)
/gene="bla"
complement(1185..2045)
CDS
/gene="bla"
/codon_start=1
/transl_table=1
/product="beta-lactamase"
/protein_id="CA05682.1"
/db_xref="GI:2623974"
/translation="MSIOHFRVALIPPEAFCLIPFAHPETIVYKRAEOLGARVY
IELDINSKILIESRPERPRPMSTFKYLQAVLSRIDGQEDLGRIRHSQNDVE
YSPYERHLIDGMVRELCSAATMSDNTANLLITIGPKELTALAHNMGHVTLE
DRMBELNEAIPNDERDTMPVAMATIRKILITDELTLASROGLIDIMMADKVAGL
LRSLPAGMFIADSGAGERSRGIIALGPDGKPSRIVIYITGSGATMDERNRQIA
ELGASLIKHW"

BASE COUNT 632 a 629 c 622 g 629 t
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2512;
Best Local Similarity 100.0%; Pred. No. 1,2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 cgcctcgtcgtttgtatgcttcatcagctcccggttcccaagatcaagcgagttaca 180
Db 1519 GCTTCGTGCTTGGATGATGATTCATTCAGCTCCGGTCCCAACGATCAAGCGAGTTACA 1578
QY 181 tbatccccaatgtctgtcaaaaaagcgttagctcccttcggtccctcgtatggtcaga 240
Db 1579 TGATCCCAATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGCTCCGATCGTTGTGAGA 1638
QY 241 agtaagttggccgcaagtgatcactcaatggtttagcagcactgtacataatctcact 300
Db 1639 AGTAAGTTGGCCGCGGTGTATCACTCATGTTATGCGACGATGCAATATTCCTTACT 1698
QY 301 gtcacgcatccgtaagaatgtcttctgtgactgtgaatcctaacaagatcattcga 360
Db 1699 GTCATCCATCCGTAAGATGCTTTTCTGTGACTGCTGATGATCTAACCAAGTCATTTCTGA 1758
QY 361 gaatagttatgcggcgacagatgtctcttcgcccgcgtcaatacaggaataacgcg 420
Db 1759 GAATAGTTATGCGGCGACGAGTTGCTTGCCTTCGCCGGGTCAATACGGGATTAATACCGCG 1818
QY 421 ccaataagcaactttaaaagtctcaatcgtgaaacgcttctcggcgcaaaatc 480
Db 1819 CCACATACGAACTTTAAAGTCTCATCATTTGAAAAACGTTCTCGGGCGAAAACTC 1878
QY 481 tcaagatcttaacgcgtgttgatcgaatcgaatcgaatcgaatcgaatcgaatcga 540
Db 1879 TCAGGATCTTACCGCTGTGTGATCCAGTTGATGATGATGATGATGATGATGATGATGATGAT 1938
QY 541 tcttaagatcttacttacttaccagcgttcttggtgtgagcaaaaaaggaagcaaat 600
Db 1939 TCTTCAGCATCTTTACTTCTTACCAAGGTTCTGTGGTGAAGCAAAAAAGCAAGCAAAAT 1998
QY 601 gcgcgaataaagggaataagggcgacacggaatgttgaatcactactcttcctttt 660
Db 1999 GCCGCAAAAAAGGGAATAAGGCGCACACGGAATGTTGAATACATACATCTCTTTT 2058
QY 661 caatatatgaacattatcagaaggttatgtgtcctcgaagggatatacatattgaatg 720
Db 2059 CAATATTTATTAACCATTTTATCAGGTTATGTTGTCTCATGAGCGGATCATATTTGAATGT 2118
QY 721 attagaataaataaacaataaggggttcgcgcacatllcccccgaataagtgcaactgac 780
Db 2119 ATTAGAATAAATAAACAATAAGGGGTTCCCGGCAATTTCCCGGAAAAAGGCCACGTGAC 2178
QY 781 gtcttaagaataatattatcatcgaacttaactataaaatagggtatcagagggccc 840
Db 2179 GTCTAAGAAACCATTTATCATCATGACATTAACCTATTAATAATAGCGGTATCACAGGCC 2238
QY 841 ttctgtcgcgcgttccgttgatgaagcgtgaaacacctcgtacacatgacgtcccgag 900
Db 2239 TTTCGTTCGGCGGTTTCGGTGTGATGACGGTGAAGAACTCTGACATGCACTCCGGGAG 2298
QY 901 acggtacacagcttctgttaagcgaatgcgcggagacagacacccgtaaggcgatca 960
Db 2299 ACGGTACACAGCTTGTCTGTAAAGCGATGCGCGAGACAGAACCCGTCAGGCGCGTCA 2358
QY 961 gcgggtgttgccgggtgttcggcggtgactgaactatgagcg 1001
Db 2359 GCGGGTGTGCGGGGTGTGCGGGGTGCTTAAGTATGCGGC 2399

RESULT 9
SYNPU8/c 2665 bp DNA circular SYN 26-JUL-1993
LOCUS SYNPU8/c
DEFINITION pUC8 cloning vector.
ACCESSION L08959
VERSION L08959.1 GI:310827
KEYWORDS
SOURCE Synthetic construct DNA.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2665)

AUTHORS Gilbert, W.
TITLE Obtained from Vecbase 3.0
JOURNAL Unpublished (1991)
COMMENT These data, and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program. pUC8 - Cloning vector

#TYPE DNA CIRCULAR

ENTRY pUC8
TITLE pUC8 - Cloning vector
DATE 17-SEP-1986
#sequence 16-DEC-1986
ACCESSION VR0021
SOURCE artificial
REFERENCE

#number 1
#authors Vieira J., Messing J.
#journal Gene (1982) 19: 259-268
#comment see also 'note added in proof'
REFERENCE

#number 2
#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
#book Cloning Vectors, Elsevier 1985 and supplements
#comment vector I-A-IV-20
COMMENT

Assembled from pUC18 and M13mp8 by F. Pfeiffer, MPI, Martinsried
Revised 16-DEC-1986 by F. Pfeiffer:
2287/8 'AT' to 'TA' to match revised sequence of pBR322
The strand shown contains the beta-galactosidase mRNA sequence
including the multiple cloning site of M13mp8.

KEYWORDS

CROSSREFERENCE

#complement
Vecbase(3):pUC8c
#parent
Vecbase(3):pUC7, Vecbase(3):M13mp8, Vecsource(3):bga18
#brother
Vecbase(3):pUC9
#offspring
Vecbase(3):pUC12, Vecbase(3):pEMBL8m, Vecbase(3):pEMBL8p,
Vecbase(3):pUC830, Vecbase(3):pOM2, Vecbase(3):pOM4,
Vecbase(3):pOM8
Vecbase(3):pUC81, Vecbase(3):pUC82, Vecbase(3):PIC7 PARENT
Features of pUC8 (2665 bp)
Residue source
1- 426 6002-6427 M13mp8
1- 229 1069-1297 Lac-Operon
230- 265 1- 36 POLYlinker of M13mp8
268- 426 1303-1461 Lac-Operon
427- 526 2351-2252 (c) pBR322
527- 663 2210-2074 (c) pBR322
664-2665 4355-2354 (c) pBR322
Conflict (cfl) and Mutations (mut):

pUC8 source
mut 1107 T C 3912 (c) pBR322
mut 1408 A G 3611 (c) pBR322
FEATURE
931-1719 1-789 Ap-R: b-lactamase
POLYLINKER EcoRI-SmaI-BamHI-SalI-PstI-HindIII
SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY pUC8 #length 2665 #checksum 2136.
Location/Qualifiers
1- 2665
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 654 a 680 c 670 g 661 t
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2665;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      1 cctgcacattatccgcctccatccagttctatatttctgcgggaagcagtaagt 60
Db      1510 CCTGCAACCTTTATCCGCCCTCCATCCAGTCTATTAATGTTGGCGGAACTAGAGTAAGT 1451
Oy      61 agtcgcagtaatatagtttcgcaacgcttgctgcatctgtctacaggaacgcgtggttca 120
Db      1450 AGTTCGCCAGTTAATAGTTTGGCAACGTTGTCATGTCATGACAGGACGTCGTGGGTCA 1391
Oy      121 cgcctgcgtcttggtatggcttcattcaagctccggttcccaagatcaaggcgattaca 180
Db      1390 CGCTCGCTGTTGGTATGGCTTCATTCACCTCCGGTTCCTCCAAACGATCAAGGCGAGTTTACA 1331
Oy      181 tgatcccccatgtttgcaaaaaaagcggttagctcctctgctccctcgtctgctgca 240
Db      1330 TGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCCTTGCTGCTCCGATCGTTGTCACA 1271
Oy      241 agtaagttgcccgaatggtatctatcactcaatggtatgagcaagcattgataatctcttact 300
Db      1270 AGTAAGTTGGCCCGCAGGTATATCAGTCATGTTATGGACAGCACTGCATTAATCTCTTACT 1211
Oy      301 gtcatagcatacgttaagatgctttctgtgactggtgagtaactcaacaaagtcattctga 360
Db      1210 GTCATGCGATCCGTAAGATGCTTTCTGTGACTGTTGAGTACTCAACCAAGTCATTTCTGA 1151
Oy      361 gaatagtatagtcgagcagatgctctctgcgcgcgtcaatacgggataataccgcg 420
Db      1150 GAATAGTGTTGGCGGCGACGAGTCTCTGCGCGCGTCATATACGGGATATATACCGCG 1091
Oy      421 ccacataagcagaactttaaagtgtcattcaatgtgaaacgcttctcgaggcgaaacac 480
Db      1090 CCACATAGCAGAACTTTAAAGTGTATCATCATGGAAGAAAGTTCCTTCGGGCGCAAAACATC 1031
Oy      481 tcaagatcttaccgctgttgagatccagtgtgaatcccaactgctgcaaccaactga 540
Db      1030 TCAAGATCTTACCGCTGTGTGAGATCCAGTTCCATGTAACCCACTGCTGACCCCACTGA 971
Oy      541 tcttcagcatcttacttacttaccacagcgttctcgtggttgagcaaaaaacaggaagcaaat 600
Db      970 TCTTCACACATCTTTACTTTCCACAGGCTTCTGGGTGAGCAAAAAAGGACGCAAAAT 911
Oy      601 gcgcgcaaaaaaggaataagggcgagcagcaggaatgttgatatactactcttctttt 660
Db      910 GCCGCAAAAAAGGAAATAAAGGGCGACACGGAAATGTTGAATCTCATACTCTTCCTTTT 851
Oy      661 caatatattgaagcatcttcaaggggtatgtctcactgaagcagatatactattgaatgt 720
Db      850 CAATATTATTGAAGCATTTATCAGGTTTATGTCTCATGAGCGGATACATATTGAATGT 791
Oy      721 attagaataataacaataaggggttcgcgcgacatctcccgaaagtgccactgac 780
Db      790 ATTAGAATAATAACAATAATAGGGGTCCCGGCACATTTCCCGAANAAGTCCACCTGAC 731
Oy      781 gtctaaagaacccattatcatgacatcttaacctataaataagggctatcagaagggcc 840
Db      730 GTCTAAGAACAACATTATTAATCATGACATTAACCTTAATAAATAGGCTATACAGAGGCC 671
Oy      841 ttctgctcgcgcttccggtgatalgacggtgaaaacctctgcacacatgacgtcccgag 900
Db      670 TTCTGCTCGCGGCTTTCGGTGAATGAGAGCGTGAAGAAACCTTGACACATGACGCTCCGAG 611
Oy      901 acggtcacagctgtctgttaagcgaatgacggtgagcagacagaccgctcagggcggtca 960
Db      610 ACGGTCAACAGCTTGTCTGTAAAGCGGATCCGCGGAGCGAGCAAGCCCGTACGGGCGGTCA 551
Oy      961 gcgggtgttgagggtgtgctggggtgctgacttaactatagcgc 1001
Db      550 GCGGGTGTGGCGGGGTGTGGGGCTGGCTTAATATATGCGGC 510

```

RESULT 10
SYNPUC9/c SYNPUC9V 2665 bp DNA circular SYN 26-JUL-1993
LOCUS pUC9 Cloning vector.


```

|||||
Db      610  ACGGTACACCTGTTGTGTAGCGGATCCGGAGACAGACAAACCCGTACAGGCGGTCA 551
Qy      961  gcgggtgttgccgggtgtcggggctgagcttaactacgcgc 1001
      |||||||
Db      550  GCGGCTTGTGCGGGGTGTCCGGGCTGAGCTTAATATGCGGC 510

RESULT  11
SYNPICTV/c LOCUS      2668 bp      DNA      circular      SYN      26-JUL-1993
DEFINITION PIC7 cloning vector.
ACCESSION  L08880
VERSION    L08880.1 GI:310785
KEYWORDS
SOURCE     Synthetic construct DNA.
ORGANISM   synthetic construct
REFERENCE  1 (bases 1 to 2668)
AUTHORS   Gilbert, W.
TITLE      Obtained from Vecbase 3.0
JOURNAL    Unpublished (1991)
COMMENT    These data and their annotation were supplied to Genbank by Will
           Gilbert under the auspices of the Genbank Curator Program. PIC7 -
           Cloning vector
           ENTRY PIC7
           TITLE PIC7 - Cloning vector
           DATE 25-FEB-1987
           ACCESSION V00090
           SOURCE artificial
REFERENCE
#number 1
#authors Marsh J.L., Erlie M., Wykes E.J.
#journal Gene (1984) 32: 481-485
#title
The PIC plasmid and phage vectors with versatile cloning
sites for recombinant selection by insertional inactivation
COMMENT
Assembled from pUC8 and Genbank:PIC7 by F. Pfeiffer
For construction of PIC7, a synthetic oligonucleotide has
been used to replace the pUC8 polylinker and thus to construct
a new cloning vector with a different polylinker. The other
PIC-vectors are based on this new PIC7 polylinker, which was
combined with the existing pUC9 and pUC19 polylinkers in the
following arrangements:
PIC19 and PICEM19 vectors:
PIC19: EcoRI - Poly (PIC7) - HindIII - Poly (PUC9) - EcoRI
PIC19H: HindIII - Poly (PUC9) - EcoRI - Poly (PIC7) - HindIII
PIC20-vectors:
PIC20: EcoRI - Poly (PIC7) - HindIII - Poly(PUC19) - EcoRI
PIC20H: HindIII - Poly(PUC19) - EcoRI - Poly (PIC7) - HindIII
COMMENT
from Genbank:
To produce greater versatility of insertional inactivation of
beta-galactosidase activity for subcloning and sequencing, a
chemically synthesized oligonucleotide, specifying nine
restriction sites including BglII, XhoI, NruI, ClaI, SacI and
EcoRV in various configurations with existing polylinkers, was
created. These improved polylinkers were inserted into plasmids
for routine cloning of ds-DNA and into chimeric phage/plasmids
for biological production of ss-DNA. The most versatile
polyrecognition pattern specifies 17 restriction sites in the
beta-galactosidase alpha-complementing gene fragment. Clone
PIC7 was used to produce all the other polylinker-carrying
vectors.
KEYWORDS
CROSSREFERENCE
#parent
Vecbase(3):pUC8, Genbank(50):PIC7
#offspring
Vecbase(3):PIC19H, Vecbase(3):PIC19,
Vecbase(3):PIC20H, Vecbase(3):PIC20,
Vecbase(3):PICEM19Hm, Vecbase(3):PICEM19Hp,

```

```

Vecbase(3):PICEM19Hm, Vecbase(3):PICEM19Hp
PARENT
Features of PIC7 (2668 bp)
residue source
1- 235 1- 235 pUC8
230- 268 1- 39 polylinker of PIC7
263-2668 260-2665 pUC8
Conflict (cfl) and Mutations (mut): none
FEATURE
934-1722 1-789 Ap-R; b-lactamase
POLYLINKER EcoRI-EcoRV-XbaI-BglII-XhoI-SacI-NruI-HindIII
SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY PIC7 #length 2668 #checksum 6564.
location/Qualifiers
1. 2668
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 658 a 677 c 668 g 665 t
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2668;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 cctgcacattatccgcctccatccagtcattatattgttcggggaagtagtaagt 60
Db      1513 CCTGCAACTTTATCCCGCTCCATCCAGTCTATTAATTTGTGCGGGAGACTGAGTAACT 1454
Qy      61 agttgccagtaataatgtttgagcaacgtttgtgcatgtccagagcagtggtgtca 120
Db      1453 AGTTGCCAGTAAATGTTTGGCAACGTTGTGTCATTGCTACAGCAATCGTGTGCTCA 1394
Qy      121 cgtcgtcgttggtatgtgcttcattcaagtcacgcgtgtcccaacagatcaaggcgttaca 180
Db      1393 CGCTGCTGTTTGATGTGGCTTATTCAGCTCCGAGTCCCAAGATCAAGCGAGTTTACA 1334
Qy      181 tgaaccacatgttgcgcaaaaaagcgltagtcctctgtgctccatgagtggtcaga 240
Db      1333 TGATCCCATGTTGCAAAAAAGCGGTAGCTCCTTGCGTCTCCGATCGTGTGTCAGA 1274
Qy      241 agtaaatgtgcgcagtgattatcatcagtgatgagcagcagctgataatctctact 300
Db      1273 AGTAAGTTGGCCGACGTTATCTCATGTTATGACAGCATGATATTTCTTTACT 1214
Qy      301 gtcatgccatccgtaagatgctttctgtgactgagtgagtaactcaaccaagtcattcga 360
Db      1213 GTCATGCCATCCGTAAGATGCTTTCTGTGACTGAGTGAATCAACCAAGTCAATTCGA 1154
Qy      361 gaatagtatgcgagcagcagtgctcttcgcccggcgtaactcaggaataatccgcg 420
Db      1153 GAATAGTATGCGGACGACGAGTGTCTTGCCCGCGCATATACGGGATATATACCGG 1094
Qy      421 ccacatagcgaacttaaaagtgctcatcatcattggaacagttcttcggggcgcaaac 480
Db      1093 CCACATAGCAGAACTTTAAAGTGCTCATCATATGGAACAGTTCTTGCGGCGCAAACTC 1034
Qy      481 tcaagatcttaccgctgttgagatccagtgatgaltlaaaccactgltgacacccaactga 540
Db      1033 TCAAGATCTTACCGCTGTGATGATCCAGTTGATGAACCCAGCTGTGACCAACTGA 974
Qy      541 tcttcagatcttacttacttaccacagcgttctggtgtgagcaaaaaaggaagcaaat 600
Db      973 TCTTACGATCTTTACTTTACACAGGCTTCTGCGGTGAGCAAAAAAGGAAGCAAAAT 914
Qy      601 gccgcaaaaaaggaataaaggcgcaacgggaatgttgaataactactcttccttt 660
Db      913 GCCGCAAAAAAGGAATAAAGGCGACAGCAAGAAATGTTGAATACTCAATCTTCTTTT 854
Qy      661 caatatattgaagcatltaacaggtatgtctcatagcgagatattgaatgt 720

```

```

Db 853 CAATATTATTGAACATTTATCAGGTTATTCATCATGAGCGGATACATATTGAATG 794
Oy 721 atttaaaaaataaacaataagggttcgcgcgcacattccccgaaaagtgcacctgac 780
Db 793 ATTGAGAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAGTGCACACTGAC 734
Oy 781 gtctaaagaacattatcatgacattaaacctaaataagcgatcacgaagcc 840
Db 733 GCTAAGAAACATTTATTCATGACATTAACCTATAAATAAGSGTATCAGAGGCC 674
Oy 841 ttctgcctgcgcgcgttcggtgatgaagcgttgaacacctgcacacatgcagctcccgag 900
Db 673 TTTCGCTCGCCGGTTCGATGATGACGGTGAAACCTCTGACACATGCAGCTCCCGGAG 614
Oy 901 acggtacagactgtctgtatgaagcgaatgcgcggaagacagaccggtcagggcggtta 960
Db 613 ACGGTACACAGCTGTCTGTAGACGGATGCCGGAGACAAAGCCGTCAGGCGCGTCA 554
Oy 961 gcgggctgtgcgcggtgtcgcggtcgtgcttaactatgcgcg 1001
Db 553 GCGGCTGTGGCGGGTGTGCGGGCTGCTTAACCTAATGCGGC 513

RESULT 12
SYNPUCTV/C 2674 bp DNA circular SYN 26-JUL-1993
LOCUS SYNPUCTV 2674 bp DNA circular SYN 26-JUL-1993
DEFINITION pUC7 cloning vector.
ACCESSION L08958
VERSION L08958.1 GI:310823
KEYWORDS
SOURCE
ORGANISM
Synthetic construct DNA.
REFERENCE
AUTHORS 1 (bases 1 to 2674)
TITLE Gilbert W.
JOURNAL Unpublished (from Vecbase 3.0)
COMMENT Unpublished (1991)
These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program. pUC7 -
ENTRY PUC7 #TYPE DNA CIRCULAR
DATE 17-SEP-1986
#sequence 16-DEC-1986
ACCESSION V80020
SOURCE artificial
REFERENCE
#number 1
#authors Vieira J., Messing J.
#journal Gene (1982) 19: 259-268
#comment see also 'note added in proof'
REFERENCE
#number 2
#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
#book Cloning Vectors, Elsevier 1985 and supplements
#comment vector I-A-IV-20
COMMENT
Assembled from pUC19 and M13mp7 by F. Pfeiffer, MPI, Martinsried
Revised 16-DEC-1986 by F. Pfeiffer:
2296/7 'AT' to 'TA' to match revised sequence of pBR322
The strand shown contains the beta-galactosidase mRNA sequence
Including the multiple cloning site of M13mp7.
KEYWORDS
CROSSREFERENCE
#complement
Vecbase(3):pUC7c
#parent
Vecbase(3):pBR322, Vecbase(3):M13mp7,
Genbank(50):Ecolac, Vecsource(3):Dgal17
#offspring
Vecbase(3):pUC8, Vecbase(3):pUC9, Vecbase(3):pUR222 PARENT
Features of pUC7 (2674 bp)
residue source

```

```

1- 435 6002-6436 M13mp7
1- 229 1069-1297 Lac-Operon
230- 277 1- 48 Polylinker of M13mp7
278- 435 1304-1461 Lac-Operon
436- 535 2351-2252 (c) pBR322
536- 672 2210-2074 (c) pBR322
673-2674 4355-2354 (c) pBR322
Conflict (cfl) and Mutations (mut):
pUC7 source
mut 1116 T C 3912 (c) pBR322
mut 1417 A G 3611 (c) pBR322
FEATURE
940-1728 1-789 Ap-R: b-lactamase
POLYLINKER EcoRI-BamHI-SalI-PstI-SalI-BamHI-EcoRI SELECTION
#resistance Ap
#indicator Beta-galactosidase
SUMMARY pUC7 #length 2674 #checksum 3876.
LOCATION/Qualifiers
1..2674
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 656 a 682 c 673 g 663 t
ORIGIN
Query Match 100.0%; Score 1001; DB 12; Length 2674;
Best Local Similarity 100.0%; Pred. No. 1, 2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 cctgcaacttatccgcctccatccatcactatcattgttgcgcggaagcagtaagt 60
Db 1519 CCTGCAACTTATCCCGCGCTCCATCCACTGATTAATTGTTGCCGGAGAACTACAGTAAGT 1460
Oy 61 agttgcgaattatgaatttgcggaacggttggtgcattgtctacaggaatcgtgtgtca 120
Db 1459 AGTTGCCAGTTAATGATTTGCCGCAAGTTGTCATTCACAGCATCGTGTGTCA 1400
Oy 121 cgcctgcgtttgtatgagcttcattcagctccggttcccaagatcaagcgaagttaac 180
Db 1399 CGCTGCTGTTGTTGATGAGCTTCACTTACGCTCCGCTCCCAAGATCAAGCGAGTTACA 1340
Oy 181 tgatccccaatgttgcgaaaaaagcgttagctccctcgttcctccgacatcgttgcaga 240
Db 1339 TGATCCCCCAATGTTGTCAAAAAGCGGTAGCTCCTTCGCTCGATCGTTGTACAGA 1280
Oy 241 agtaagttgcccgcgaggttatcactcattatgttatgacagcactgtcataatttctact 300
Db 1279 AGTAAGTTGGCCGAGGTTATCATCATCATGTTATGCGACACATGATTAATTCCTTACT 1220
Oy 301 gtcataccatccgttaagatgctttctgtgactgcgtgcgtgagtaactcaacaaatcattcga 360
Db 1219 GTCATGCCATCCGTAAGATGCTTTTCTGTACATGCTGAGTACTCAACCAAGTATTTCTGA 1160
Oy 361 gaatagtgtatgcgcgacccagatgctcttgcgcgcggtlcaatacggataacacgcg 420
Db 1159 GAATAGTGTATGCGGCGACGAGTGTCTTGTCCCGGCGCATATGCGGATATACCGGG 1100
Oy 421 ccacatagcgaacttaaaagtgctcattcattatggaagaacgttcttcggggcgaaacc 480
Db 1099 CCACATAGCAGAACTTTAAAGTGCTCATCATTAATGGAAGAGTTCTTGGGGCGAAAAAC 1040
Oy 481 tcaagatcttacgcgtgttgagatccagttcagatgaacccaactgcgtgacccaactga 540
Db 1039 TCAAGGATCTTACCGCTGTGAGATGCAGTTCATTAACCACTGTGTACCCCAACTGA 980
Oy 541 tcttcagatccttactccttaccagcgttctcgtgtgcgaacaaacaggaagcaaat 600
Db 979 TCTTCAGCATCTTTACTTTACACACAGCTTTCGGGTGAGCAAAAAACAGAGCAAAAT 920
Oy 601 gccgcaaaaaaggaataaggcgacacggaataatgttgatactcattcttcttctt 660
Db 919 GCGCAAAAAAGGAATTAAGGGCGACGCAAAATGTTGAATCACTACTCTTCTTTT 860

```

QY	661	caatattatgaacatttttcaaggattttctcatalagacgaataacatttgaat	720
Db	859	CAATATTATTGAACCATTTATCAGAGGTTTATGTCTCATGAGCGGATTCATATTTGAATGT	800
QY	721	atttagaaaaataacaataataggggtttccgcgcacatttcccgaaaaagtgcacctgac	780
Db	799	ATTTAGAAAAATTAAACAATATGAGGGGTTCCGGCAGCATTTCCCGAAAAGTGCACCTGAC	740
QY	781	gtctaaagaacattatattatcatfacattaacctataanaatagcgtalacagagccc	840
Db	739	GTCATAGAAACCATTTATTATCATGACATTAACCTATAAATAATGAGGCTATACAGAGGCC	680
QY	841	tttcgctctcgcggtttccggttgatgacagcgtgaaaaaccttcgaacatgacgctccggag	900
Db	679	TTTTCGCTCTCGGCGCTTTCCGGTGATGACAGCGTGAACAACTTCGACACATGACAGCTCCGGAG	620
QY	901	acggtcaaacagctttctgttaagcgcgtatgcgcggagacgaagaacccgtcaagggcgctca	960
Db	619	ACGGTCACAGCTTCTCTGTATAGCCGATATGCCGGAGCAGACAGCCCTTCACAGGCGCTCA	560
QY	961	gcggcgtgtgcgcgggtgtcgcagcgtgcgttaactaatatgcgcg	1001
Db	559	GCGGCTGTGGCGGGGTGTCGGGCTGTGGCTTAACATATCGGCG	519

RESULT	13		
A02712/c			
LOCUS	A02712	2680 bp	DNA
DEFINITION	pUC13 DNA sequence.		PAT
ACCESSION	A02712		27-APR-1993
VERSION	A02712.1		
KEYWORDS	GI:344656		
SOURCE	.		
ORGANISM	synthetic construct.		
REFERENCE	synthetic construct.		
	artificial sequence.		
	1 (bases 1 to 2680)		

PROKARYOTES
Patent: WO 8809373-A 10 01-DEC-1988
JOURNAL
FEATURES
source
location/Qualifiers
1. .2680
Location: "Mathematical constants"

```

/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT      657 a      683 c      676 g      664 t
ORIGIN

```

Query Match	100.0%;	Score 1001;	DB 6;	Length 2680;
Best Local Similarity	100.0%;	Pred. No. 1.2e-274;		
Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	ccccaacttatccgctccatccagctcattaatgctgttcgcgggaagctcagaagtaagt	60
Db	1525	cttgcaaatattatccgctccatccatccattattattgtttgccgggaagctcagaagtaagt	1466
OY	61	attccgcagtaataaatttttcgcaacgcttbtgtccattgtctacagcaatcgtagtgc	120
Db	1465	agttcgccagattaatagtgttggcccaacggtgttgccattgtctacagcaatcgtagtgc	1406
OY	121	cgcctgcgtttgttatggcttcatcgaatccgcggtctcccaacgataaagcgaaattca	180
Db	1405	cgcctgcgtttgttatggcttcatcgaatccgcggtctcccaacgataaagcgaaattca	1346
OY	181	tgaatccccatgttgcacaanaagcggttaagctccctgcgtccatccgattgtcaga	240
Db	1345	tgaatccccatgttgcacaanaagcggttaagctccctgcgtccatccgattgtcaga	1286
OY	241	agtaagttgscgcgaagttatcaactcaatgattatagtcgaactgaataattcttaact	300
Db	1285	agtaagttgscgcgaagttatcaactcaatgattatagtcgaactgaataattcttaact	1226
OY	301	gtcatgcatactcgaagaatccttctctgtgaactcgtgtgagtaactcaacaagtaattctcga	360

Db	1225	GTCAATGCATCCGTAAGATGCTTTCTGTACATGATGATGACTCAACCAAGTCAATCTGA	116
OY	361	gaataatgatatgagcgacccgaagtctctcttgcgccgagctcaataagagataataccgg	420
Db	1165	GAATAGTGTATGCGCGCACCGAGATGCTCTTGGCCCCGGCTCAATACGGGATTAATCCGGC	110
OY	421	cccatatgcaaaactttaaagctcatcatatgaaacgctcttcctggcgcaaaactc	480
Db	1105	CCCAATACCAAGACTTTTAAAGTCTCATCTTGGAAAGAGTTCTTGGGGCCAAAACTC	104
OY	481	tcaagatctaccgcgctgttgaatccagttcgaatgtaaacccactcgtgcaccaactga	540
Db	1045	TCAAGATCTTACCGCTGTTGATGATCCAGTTCGATGTAAACCCACTCGTGCACCCAACTGA	986
OY	541	tcttcagcatctttacttttaccagcgctttctgggtagcaaaaagaagaagcaaat	600
Db	985	TCTTACGATCTTTTACTCTTCACGAGGTTCTGGGTGACCAAAACAGGAAGGCANAAT	926
OY	601	gcgcgaanaaagggaataaaggcgacacggaatgttgaatactacatccttcctttt	660
Db	925	GCCGCAAAAAGGGAATAGGGCGCACACGAAAGTGGATACATCACTTCCTCTTTT	866
OY	661	caatatattgaagcatttalcagaggtatgtctcatalgacgagatacatatttgaatc	720
Db	865	CAATATTATTGAGCATTTATCAGAGGTTATGTCTCATGACCGCATTCATATTTTGAATGT	806
OY	721	atttagaaaataaacaatatagggtctccgcgcacatttcccgaaaagcgcaactgac	780
Db	805	ATTTAGAAAATAAACAATATGAGGGGTTCCGGCATTTTCCCGAAAAGTGCCACTGAC	746
OY	781	gtctaaagaacatattatcatcatcaatlaaacataaaataagcgctatcacagagccc	840
Db	745	GCTTAGAAGAACCATTTATTATCATGACANTTAACCTATAAATAGGCTATACAGAGGCC	686
OY	841	tttcgcttcgcgcgcttcgctgatatgcggtgaaaaacctctgacacatgcagctcccgag	900
Db	685	TTTCTCTCTCGCGCTTGTGGTGTAGACGGTGAAGAAACCTCTGCACATGCAAGCTCCCGAG	626
OY	901	acagtcacaagcttgcgtctaaagcgagatgcgcggagagcgaaagccgctaaaggcgctca	966
Db	625	ACGGTCACAGCTTGCTGTETAAGCGGATGCCGGAGCAACAGGCCCTTCAGGCGCCTCA	566
OY	961	gcggatgttgcgcggtgctcgaggctcgtgcttlaactacatgcgcg 1001	
Db	565	GCGGCTGTGGCGGGGTGTCGGGGGCTGCTTAACTATGCGGC 525	

RESULT 14				
LOCUS	SYNPUC12V/c			
DEFINITION	SYNPUC12V 2680 bp	DNA	circular	26-JUL-1993
ACCESSION	pUC12 cloning vector.			
VERSION	L09129			
KEYWORDS	L09129.1 GI:310821			
SOURCE				
ORGANISM	Synthetic construct DNA.			
REFERENCE	Synthetic construct			
AUTHORS	artificial sequence.			
TITLE	1 (bases 1 to 2680)			
JOURNAL	Gilbert,W.			
COMMENT	Obtained from VecBase 3.0			
	Unpublished (1991)			
	These data and their annotation were supplied to GenBank by Will			
	Gilbert under the auspices of the GenBank Curator Program			
	pUC12 -			

Cloning vector	#TYPE	DNA	CIRCULAR
ENTRY	PUC12		
TITLE	PUC12 - Cloning vector		
DATE	17-SEP-1986		
#sequence	16-DEC-1986		
ACCESSION	V80023		
SOURCE	artificial		
REFERENCE			
#number	1		

#authors Vieira J., Messing J.
#journal Gene (1982) 19: 259-268
#comment see 'note added in proof'

REFERENCE

#number 2
#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
#book Cloning Vectors, Elsevier 1985 and supplements
#comment vector I-A-iv-20

COMMENT

Assembled from pUC18 and M13mp10 by F. Pfeiffer, MPI,
Marlinsried

Revised 16-DEC-1986 by F. Pfeiffer:

2302/3 'Ap' to 'TA' to match revised sequence of PBR322

The strand shown contains the beta-galactosidase mRNA sequence
including the multiple cloning site of M13mp10. KEYWORDS

CROSSREFERENCE

#complement

VecBase(3):pUC12c

#parent VecBase(3):pUC8, VecBase(3):M13mp10, VecSource(3):bca112

#brother VecBase(3):pUC13

#offspring

VecBase(3):pUC18, VecBase(3):pSP64, VecBase(3):pCEM1,
VecBase(3):pSP6T3, VecBase(3):pT712, VecBase(3):pCKSP6 PARENT

Features of pUC12 (2680 bp)

residue source

1-441 6006-6442 M13mp10

1-229 1069-1297 Lac-Operon

230-280 1-51 polylinker of M13mp10

283-441 1303-1461 Lac-Operon

442-541 2351-2252 (c) PBR322

542-678 2210-2074 (c) PBR322

679-2680 4355-2354 (c) PBR322

Conflict (cfl) and Mutations (mut):

pUC12 source

mut 1122 T C 3912 (c) PBR322

mut 1423 A G 3611 (c) PBR322

FEATURE

946-1734 1-789 Ap-R: b-Lactamase
POLYLINKER EcoRI-SacI-SmaI-BamHI-XbaI-SalI-PstI-HindIII SELECTION

#resistance Ap

#indicator beta-galactosidase

SUMMARY pUC12 #length 2680 #checksum 9725.

location/Qualifiers

1..2680

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 657 a 685 c 674 g 664 t

ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2680;

Best Local Similarity 100.0%; Pred. No. 1.2e-274;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccgcgaacttaccgcccacatccatctataattgttcgcggaagcctagagtaagt 60
Db 1525 CCTGCACACTTATCCGCCCATCCAGTCTATTATGTGCGGGAAGCTAGATAAGT 1466
QY 61 agtcgcagtaataagtttcgcgaagttgttcgcatgtctacagagcagtcgtgtgca 120
Db 1465 AGTTGGCCAGTTAATAGTTGGCAACGTTGTTGCCATGTGTACAGGCGATGCTGTGCA 1406
QY 121 cgcctcgttctgtatagcttcattcagctcgcgttcccaagatcaagcaggttaca 180
Db 1405 CCCTCGCTCTTGTGTATGCTTCATTCAAGCTCCGGTCCCAACGATCAAGCGAGTATACA 1346
QY 181 tgaatcccaattgttgcgaagaacggttagctctcctcgtgcgcctccagatggttgcaga 240
Db 1345 TGAATCCCATCTTGTGCAAAAGCGGTGTAGCTCTTCCGCTCCCATGCTGTGTGAGA 1286
QY 241 agtaagttggccgagttatcatcactagtgtaagcagcactgacataatctctact 300

Db 1285 AGTAAGTTGGCCGCGAGTGTATCACTCATGTGTAAGCAGCAGCATCAATTCCTTACT 1226
QY 301 gtcatgcatccgttaagatgtcttctgtgactggtgtgacttaaccaagatcttctga 360
Db 1225 GTCAATGCGATCCGTAAAGATGCTTTCTGTGACTGGTGTACTCAACCAAGTCAATTCGA 1166
QY 361 gaatagttatagc 420
Db 1165 GAATAGTGTATGCGCGCAGCAGCAGTGTGCTTGTCCCGGCGTCAATACGGGATATACCGCG 1106
QY 421 ccacatagcagaacttaaaagtgtcatcatlgtgaaaacgttcttcggygcgaaaatc 480
Db 1105 CCACATFACAGAACTTTAAAGTGCATCAFTTGAAAAAGCTTCTGCGGCGCAAACTC 1046
QY 481 tcaagatcttaacgcgtgtttagatccaagttcgatgtaaccactgtgacccaactga 540
Db 1045 TCAAGATCTTACCCGCTGTGATCCAGTTGATGTAACCCACTGTCACCCACTGA 986
QY 541 tcttaagatcttacttcttaccagcgttctgtgtgtaagcaaaaacgaagacaaat 600
Db 985 TCTTAGCATCTTTACTTTACTTTCACACAGCGTTTCTGGGTAGCAAAAACAGAGCAAAAT 926
QY 601 gccgcaaaaaggaataaaggcgacacggaatgttgaatactatactctctctttt 660
Db 925 GCCGCAAAAAGGCAATTAAGGCGCACAGCAAAAGTTGAATATCATATCTTCTTTT 866
QY 661 caataatttaagcattatcatcaggttatgtctctcttgagcggatcatatatgtatgt 720
Db 865 CAATATTAATGAAGCATTTATACAGGTTATTTGTTCTCATGACCGGATCAATATTTGAATGT 806
QY 721 attagaanaataaacaataaggggttcgcgcacatltccgcgaagaagtgcacgtgac 780
Db 805 ATTATGAATAAATAAATAATAGGGGTTCCGCGCATTTTCCCGCAAAAGTCCACCTGAC 746
QY 781 gtctaagaacattatatacatgaacattaaactataaataagcgtatcacagggccc 840
Db 745 GCTTAAGAAACCATTAATATCATGATGATTAACCTATTAATAATAGCGGTATCACAGGCGCC 686
QY 841 tttgtctcgcgcgcttctcgtgtagatgacggtgtaaaactctgacacatgcagctccggag 900
Db 685 TTTGCTCTCGCGCTTTCGCTGATGACGCTGTAATAAACCCTGTACATGACAGCTCCCGAG 626
QY 901 acggtcacagcttctcgttaagcagatgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 960
Db 625 ACGGTCAACGCTTGTCTGTATAGCGGATGCGCGAGACAGACAAAGCCCTCAGGGCGGTCA 566
QY 961 gcgggtgttgccgggtgtcgcgggcttgcttaactatgcgcgc 1001
Db 565 GCGGATGTTGGCGGCTGCGGCGCTGAGCTTAACATATCGCGC 525

RESULT 15

SYNPOC13V/c

LOCUS SYNPOC13V 2680 bp DNA circular SYN 06-MAR-2001

DEFINITION Cloning vector pUC13, complete sequence.

ACCESSION L09130

VERSION L09130.1 GI:310822

KEYWORDS

SOURCE

ORGANISM

Cloning vector pUC13.

Cloning vector pUC13

artificial sequence; vectors.

REFERENCE

1 (bases 1 to 2680)

AUTHORS

Vieira, J. and Messing, J.

TITLE

The pUC plasmids, an M13mp7-derived system for insertion

mutagenesis and sequencing with synthetic universal primers

Gene 19 (3): 259-268 (1982)

JOURNAL

MEDLINE

6295879

PUBMED

2 (bases 1 to 2680)

REFERENCE

Gilbert, W.

Obtained from Vecbase 3.0

JOURNAL

Unpublished

COMMENT

These data and their annotation were supplied to Genbank by Will Gilbert under the auspices of the GenBank Curator Program. pUC13 - Cloning vector

#TYPE DNA CIRCULAR

ENTRY pUC13
TITLE pUC13 - Cloning vector

DATE 17-SEP-1986

#sequence 16-DEC-1986

ACCESSION V80024

SOURCE artificial

REFERENCE

#number 1

#authors Vieira J., Messing J.

#journal Gene (1982) 19: 259-268

#comment see 'note added in proof'

REFERENCE

#number 2

#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.

#book Cloning Vectors, Elsevier 1985 and supplements

#comment vector I-A-iv-20

COMMENT

Assembled from pUC19 and M13mp11 by F. Pfeiffer, MPI, Martinsried

Revised 16-DEC-1986 by F. Pfeiffer:

2302/3 'AT' to 'TA' to match revised sequence of pBR322

The strand shown contains the beta-galactosidase mRNA sequence including the multiple cloning site of M13mp11. KEYWORDS

CROSSREFERENCE

#complement

Vecbase(3):pUC13c

#parent

Vecbase(3):pUC9, Vecbase(3):M13mp11, Vecsource(3):bgal13

#brother

Vecbase(3):pUC12

#ofspring

Vecbase(3):pUC19, Vecbase(3):pUC19c,

Vecbase(3):pSP65, Vecbase(3):pSEM2, Vecbase(3):pT713 PARENT

Features of pUC13 (2680 bp)

residue source

1-441 6002-6442 M13mp11

1-230 1069-1298 Lac-Operon

233-283 1-31 polylinker of M13mp11

284-441 1304-1461 Lac-Operon

442-541 2351-2252 (c) pBR322

542-678 2210-2074 (c) pBR322

679-2680 4355-2354 (c) pBR322

Conflict (cfl) and Mutations (mut):

pUC13 source

mut 1122 T C 3912 (c) pBR322

mut 1423 A G 3611 (c) pBR322

FEATURE

946-1734 1-789 Ap-R; b-lactamase

POLYLINKER HindIII-PstI-SalI-XbaI-BamHI-SmaI-EcoRI SELECTION

#resistance Ap

#indicator Beta-galactosidase

SUMMARY pUC13 #length 2680 #checksum 217.

location/Qualifiers

1..2680

/organism="Cloning vector pUC13"

/db_xref="taxon:134051"

1..441

/note="M13mp11"

1..230

/note="Lac-Operon"

233..283

/note="M13mp11 polylinker"

284..441

/note="Lac-Operon"

442..2680

/note="pBR322"

879..1739

/note="Ap-R"

/codon_start=1

/transl_table=11

/product="beta-lactamase"
/protein_id="AAK16201.1"
/db_xref="GI:13336870"
/translation="MSIOHRYVALIPFEAFCLPVEAFPELVKYNDAEDLGARVY
IELDNSGKILSEFRPEPRPMSTFKYLCAQVLSRIDAGEOLGRRVHSQNDLVE
YSPVTEKHLTDGNTVRELSAAITMSDNTANILLITIGPKELTFLNNKDDHYTRL
DRWEPELNEATFMDERDTTPVAMATTLKLLTGELTLTASRQQLVAMEAUKVAGPL
LRSAIPAGWFIADKSGAGERSGIIIALGPDGKPSRIVIVYITTSQATMDERNROIA
EIGASLIKHW"
variation
1122
/replace="c"
variation
1423
/replace="g"
BASE COUNT 657 a 683 c 676 g 664 t
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2680;
Best Local Similarity 100.0%; Pred. No. 1,2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacattatccgcctccatccagtcattatgttgcgggaagctagaatgaat 60
Db 1525 CCGCACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAAGCTAGAGTAAGT 1466
QY 61 agtcgcagtaataagttgcgaacgctgttgcattgtctacagcagcgtggtgca 120
Db 1465 AGTTCGCGCACTTAATAGTTTGGCAAGCTTGGCATTCCTACAGGCAATCGTGTGTC 1406
QY 121 cgttcgtcgttgtatgtgttcattcagtcctcgttcccaagatccaagcgagttaca 180
Db 1405 CGCTCGTGTGTTGATAGGCTTCATTCAGCTCGGTTCCCAAGATCAAGCGAGTTTAA 1346
QY 181 tgaatcccaatgttgcgaacaaagcggttaagtccttcgttccttcgttcgttcaga 240
Db 1345 TGAATCCCAATGTGTTGCAAAAAAGCGGTAGTCTCTGCGTCTCGCATCGTTGTCACA 1286
QY 241 agtaagttgcgcgaagtgatatacctcaatggttaagcagcactgcatatcttact 300
Db 1285 AGTAAGTTGGCCGACAGTGTATCATCATGTTATGGACACTGCATTAATCTCTTACT 1226
QY 301 gtcattgccatccgtaagatgcttctctgtgaactggtgagttactccaacgaatcttca 360
Db 1225 GTATGCCATCCCTTAAGATCCTTTCTGTACGTCGTCAGTACTCAACCAAGTATTTCTGA 1166
QY 361 gaatagttatgcgcgacagtgctctctgcgcgcgtccaatacgggaataaccgcg 420
Db 1165 GAATAGTGTATGCGCGACCGAGCTGCTCTGCGCGGCTCAATACGGGATATACCGCG 1106
QY 421 ccacatagaagaacttaaaagtgctcatcatcttggaacacgttcttcggggcgaaacac 480
Db 1105 CCACATAGCAAGACCTTAAAGTGCATCATCTTGAAGAAAGCTTCTTGGGGGAAAAAC 1046
QY 481 tcaagatcttaccgctgttgagatccagttcgatgtaaccacactggtcacccaactga 540
Db 1045 TCAAGATCTTACCGCTGTTGAGATCCAGTTCATTAACCACTGTGCACCCAACTGA 986
QY 541 tcttcagcatcttacttccaccagcgttctcgttggaacaaacaggaagcgaat 600
Db 541 tcttcagcatcttacttccaccagcgttctcgttggaacaaacaggaagcgaat 600
QY 541 tcttcagcatcttacttccaccagcgttctcgttggaacaaacaggaagcgaat 600
Db 541 tcttcagcatcttacttccaccagcgttctcgttggaacaaacaggaagcgaat 600
QY 601 gccgcaaaaaggaataagggcgagacaggaatggtgaatactctactctctctt 660
Db 601 gccgcaaaaaggaataagggcgagacaggaatggtgaatactctactctctctt 660
QY 661 caatatattgaagcattatcaggggtatgtctcatagaaggatatacttgaatgt 720
Db 661 caatatattgaagcattatcaggggtatgtctcatagaaggatatacttgaatgt 720
QY 721 attagaataaataaacaataaggggttccgcgacatcttcccgaaagatgccaactgac 780
Db 721 attagaataaataaacaataaggggttccgcgacatcttcccgaaagatgccaactgac 780
QY 805 atttagaaaaataaacaataaggggttccgcgacatcttcccgaaagatgccaactgac 746
Db 805 atttagaaaaataaacaataaggggttccgcgacatcttcccgaaagatgccaactgac 746

QY 781 gtctaagaacattatattatcatgacattaaacctaataaataaggcfratcacgagggcc 840
|||||
Db 745 GTCCTAAGAAACCATTTATTTATTCATGACATTAACTATAAAATAGCGGTATCACAGAGGCC 686
|||||
QY 841 ttctgtctcgcgcgttttcgltgatgacggtgaaaccltgacacatlgcaagctcccgag 900
|||||
Db 685 TTTCCGTCCTCGCGCGTTTCGGGTATGACGCGTGAACCTCTGACACATGACAGCTCCCGAG 626
|||||
QY 901 acggtcacagcttgcgtgaagcggaatgcgggagcaagaagcccgatcaaggcggtca 960
|||||
Db 625 ACGGTACACAGCTTGTCTGTAAAGCGGATGCCGAGACAGCAAGCCCGTACAGGCGCGTCA 566
|||||
QY 961 gcgggtgttgcggtgtgcgggctggtcttaactatgcggc 1001
|||||
Db 565 GCGGGGTGTGGCGGCTGCGGCTGGCTTAACATAAGCGGC 525
|||||

Search completed: January 17, 2002, 16:24:34
Job time: 13795 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 17, 2002, 16:32:59 ; Search time 343.96 Seconds
(without alignments)
2495.007 Million cell updates/sec

Title: US-09-740-211-13_COPY_6000_7000

Perfect score: 1001
Sequence: 1 cctgcacttaccgcgcgc.....ggtcgttaactatgcgcgc 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_1101: *
1: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1001	100.0	2462	21	AAA74638
2	1001	100.0	2686	22	AAAF59053
3	1001	100.0	3190	19	AAV32980
4	1001	100.0	3253	19	AAV34339
5	1001	100.0	3331	21	AAZ9245
6	1001	100.0	3343	11	AAQ04655
7	1001	100.0	3383	21	AAAS1632
8	1001	100.0	3444	21	AAAS1630
9	1001	100.0	3819	18	AAAT8825
10	1001	100.0	3819	19	AAV39266
11	1001	100.0	3819	20	AAZ22020

12	1001	100.0	3858	22	AAAD0947	Plasmid PRK50 used
13	1001	100.0	3880	19	AAV39242	Plasmid pCK7-96 nu
14	1001	100.0	3881	18	AAAT78801	Kappa light chain
15	1001	100.0	3881	20	AAZ21996	Nucleotide sequenc
16	1001	100.0	3903	22	AAAF26078	Feline IL-16 encod
17	1001	100.0	3938	20	AAAF61041	Plasmid pFRA6-Kanx
18	1001	100.0	4009	19	AAV00679	Plasmid pRHA43 con
19	1001	100.0	4045	15	AAO70942	Plasmid pND211 con
20	1001	100.0	4074	21	AAZ48267	Plasmid pT3NAM1 co
21	1001	100.0	4118	18	AAAT69188	Construct pGEM-hTR
22	1001	100.0	4118	18	AAAT69189	Construct pGEM-hTR
23	1001	100.0	4189	21	AAAF51634	Plasmid pGM712 con
24	1001	100.0	4201	22	AAAF26077	Feline IL-12p35 en
25	1001	100.0	4421	17	AAAT58319	DNA sequence of be
26	1001	100.0	4522	22	AAAF26076	Plasmid pK01-f1112
27	1001	100.0	4522	22	AAAF26094	altR reading frame
28	1001	100.0	4554	21	AAAC55541	Plasmid pSUN387
29	1001	100.0	4603	11	AAO04010	Plasmid pUCM5-6h73
30	1001	100.0	4613	22	AAAF59062	DNA sequence of p1
31	1001	100.0	4618	21	AAZ39628	Nucleotide sequenc
32	1001	100.0	4694	22	AAAF55225	Murine IAD alpha c
33	1001	100.0	4713	19	AAAV12067	Gamma heavy chain
34	1001	100.0	4723	18	AAAT78802	Plasmid pCG7-96 nu
35	1001	100.0	4723	19	AAAV39243	Nucleotide sequenc
36	1001	100.0	4723	20	AAZ21997	Nucleotide sequenc
37	1001	100.0	4723	22	AAAF55224	Murine IAD beta ch
38	1001	100.0	4724	19	AAAV12068	Expression constru
39	1001	100.0	4776	20	AAAT7617	Plasmid pHC65 nuc
40	1001	100.0	4776	20	AAAT7614	Nucleotide sequenc
41	1001	100.0	4926	19	AAV39291	Retroviral vector
42	1001	100.0	4926	20	AAZ22045	Transdominant effe
43	1001	100.0	4950	18	AAV03801	pcMV-I-Cre-PA vect
44	1001	100.0	4950	22	AAAC82936	
45	1001	100.0	4960	22	AAAD09269	

ALIGNMENTS

RESULT 1
ID AAA74638 standard; DNA: 2462 BP.
XX AAA74638;
AC
XX
XX
05-DEC-2000 (first entry)
XX
XX
Plasmid pSP72.
DE
XX
XX
Plasmid pSP72; NSF-p25 transgene; p25; cdk5 activator;
KW neuron specific enolase; NSF; promoter; neurodegenerative disease;
KW Alzheimer's disease; Parkinson's disease; amyloid lateral sclerosis;
KW Huntington's disease; traumatic brain injury; stroke; transgenic animal;
KW spinocerebellar degeneration; tau hyperphosphorylation; ds.
XX
XX
Escherichia coli.
OS
XX
XX
EP1026251-A2.
PN
XX
XX
09-AUG-2000.
PD
XX
XX
02-FEB-2000; 2000EP-0300797.
PF
XX
XX
03-FEB-1999; 9905-0118478.
PR
XX
XX
(PFIZ) PFIZER PROD INC.
PA
XX
XX
Ahlijanian MK, Mcneish JD;
PI
XX
XX
WPI; 2000-507252/46.
DR
XX
XX
New recombinant DNA molecule comprising gene encoding p25, useful for
PT producing transgenic animal to serve as model for

PT Alzheimer's-associated neurofibrillary tangle formation -
 XX Disclosure: Page 16-17; 32pp; English.
 XX
 CC The present sequence is the commercially available plasmid pSP72. The
 CC SV40 polyadenylation sequence, the rat neuron specific enolase (NSE)
 CC promoter and the human cDNA for p25 were cloned into plasmid pSP72 to
 CC generate the NSE-p25 transgene. This was used in the production of
 CC transgenic mice overexpressing human p25, an activator of protein kinase
 CC cdk5. Overexpression of p25 is sufficient to produce hyperphosphorylation
 CC of tau, which is seen in the neurofibrillary tangles associated with
 CC Alzheimer's disease. The transgenic animals are useful as in vivo systems
 CC for screening potential therapeutic compounds for their ability to
 CC inhibit or prevent the production of hyperphosphorylated tau and
 CC associated neuronal death. The transgenic animals are suitable for use as
 CC disease models of neurodegenerative diseases and tau-related pathologies,
 CC such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral
 CC sclerosis, Huntington's disease, stroke, traumatic brain injury,
 CC spinocerebellar degeneration. They are also useful as models of cdk5/p25
 CC and tau biochemistry, and to establish the role of the human p25 in the
 CC formation of hyperphosphorylated tau in neurodegenerative conditions.
 CC
 SO Sequence 2462 BP; 623 A; 614 C; 606 G; 619 T; 0 other;

Query Match 100.0%; Score 1001; DB 21; Length 2462;
 Best Local Similarity 100.0%; Pred. No. 3.4e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttattcgcctccatccagcttctattatgttgcgggaagctagaaagt 60
 DB 1349 cctgcaacttattcgcctccatccagcttctattatgttgcgggaagctagaaagt 1408
 QY 61 agttgcgcagtaaatagtttgcgaacgttgccttgcattgttcgaagagatcgtgtgtca 120
 DB 1409 agttgcgcagtaaatagtttgcgaacgttgccttgcattgttcgaagagatcgtgtgtca 1468
 QY 121 cgtctcgttctgtatgtatgtatcattcagctcgtgtcccaagatcaagcgaaattca 180
 DB 1469 cgtctcgttctgtatgtatgtatcattcagctcgtgtcccaagatcaagcgaaattca 1528
 QY 181 tttatcccatcttctgtgcaaaaaaagcgtttagctccctcgttcccgatcgtgttca 240
 DB 1529 tttatcccatcttctgtgcaaaaaaagcgtttagctccctcgttcccgatcgtgttca 1588
 QY 241 agtaagtttgcgcagtggtatcactcatgtatgtacgacacgtcacaattccttact 300
 DB 1589 agtaagtttgcgcagtggtatcactcatgtatgtacgacacgtcacaattccttact 1648
 QY 301 gtcattgcacatccgtaagatgtcttctgttactgtgtgtaactcaaccaatcattctga 360
 DB 1649 gtcattgcacatccgtaagatgtcttctgttactgtgtgtaactcaaccaatcattctga 1708
 QY 361 gaatagtttgcgcagcgcagtggtctgtcgtccggttcaatacggataatccgcg 420
 DB 1709 gaatagtttgcgcagcgcagtggtctgtcgtccggttcaatacggataatccgcg 1768
 QY 421 ccacataagcagaacttaaaaatgctcalttgaaaaagcttctcggggcgaaaactc 480
 DB 1769 ccacataagcagaacttaaaaatgctcalttgaaaaagcttctcggggcgaaaactc 1828
 QY 481 tcaagatcttaacgcgtcttgagatccagttcogatgttaaccactcgttgcacccaactga 540
 DB 1829 tcaagatcttaacgcgtcttgagatccagttcogatgttaaccactcgttgcacccaactga 1888
 QY 541 tcttagatcttcttacttaccacagcgttctcgtgtgtagcaaaaaacggagcaaat 600
 DB 1889 tcttagatcttcttacttaccacagcgttctcgtgtgtagcaaaaaacggagcaaat 1948
 QY 601 gccgcaaaaaagggaataaggcgacacgaaatgttgaatacctacatacttcttcttt 660
 DB 1949 gccgcaaaaaagggaataaggcgacacgaaatgttgaatacctacatacttcttcttt 2008

QY 661 caatatatgaagcattatccaggttatgtctctatgagcggatcatcatattgaatgt 720
 DB 2009 caatatatgaagcattatccaggttatgtctctatgagcggatcatcatattgaatgt 2068
 QY 721 attagaanaataaacaatagaggttccgcgcacatttcccgcaaaaagtccaccgac 780
 DB 2069 attagaanaataaacaatagaggttccgcgcacatttcccgcaaaaagtccaccgac 2128
 QY 781 gtctaaagaacattatattatcatatgacattaccataaaatagcgctatcagaagccc 840
 DB 2129 gtctaaagaacattatattatcatatgacattaccataaaatagcgctatcagaagccc 2188
 QY 841 ttctgcgcgcggttctgcgtatgacggtgaacacctgcacacatgcagctcccgag 900
 DB 2189 ttctgcgcgcggttctgcgtatgacggtgaacacctgcacacatgcagctcccgag 2248
 QY 901 aacgtacacagcttctgtatgaagcgatgcggagacagacaaagccgtcaggcgctca 960
 DB 2249 aacgtacacagcttctgtatgaagcgatgcggagacagacaaagccgtcaggcgctca 2308
 QY 961 gccggtgtgtgcgggtgtcgcggggtcgtgcttaactatgcgc 1001
 DB 2309 gccggtgtgtgcgggtgtcgcggggtcgtgcttaactatgcgc 2349

RESULT 2

AAF59053/C
 ID AAF59053 standard; DNA; 2686 BP.

AAF59053;
 AC

23-APR-2001 (first entry)
 DT

Plasmid vector pUC18M5 nucleotide sequence SEQ ID NO:70.
 DE

Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KW

T cell epitope; Antisugipollinosis; ds.
 KM

Cryptomeria japonica.
 OS

Synthetic.
 OS

JP2000327699-A.
 PN

28-NOV-2000.
 PD

15-MAR-2000; 2000JP-0071710.
 PF

15-MAR-1999; 99JP-0068316.
 PR

(HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA

(SANY) SANKYO CO LTD.
 PA

WPI; 2001-185061/19.
 DR

Novel peptide and its use -
 PT

Example 11; Page 61-63; 75pp; Japanese.
 PS

The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (1) has a formula of:
 CC

where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an antisugipollinosis agent.
 CC

AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.
 CC

Sequence 2686 BP; 661 A; 685 C; 675 G; 665 T; 0 other;
 SO

Query Match 100.0%; Score 1001; DB 22; Length 2686;
 Best Local Similarity 100.0%; Pred. No. 3.5e-263;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacttatacgcctccatccatcacttaattgttcggaagctagaagt 60
DB 1531 cctgaacttatacgcctccatccatcacttaattgttcggaagctagaagt 1472
QY 61 agtcgcagtaataagtttcgcaagctgttcacattgtaagcagtcgtgtc 120
DB 1471 agtcgcagtaataagtttcgcaagctgttcacattgtaagcagtcgtgtc 1412
QY 121 cgtctcgtttgttgaatgcttcacatccgctcccaagcaagcagttaca 180
DB 1411 cgtctcgtttgttgaatgcttcacatccgctcccaagcaagcagttaca 1352
QY 181 tgatcccccatgtgtgcaaaaagcgtttagctccttcggttcctcgtatc 240
DB 1351 tgatcccccatgtgtgcaaaaagcgtttagctccttcggttcctcgtatc 1292
QY 241 agtaagttgcccgcagctgtatcactcgtttagcagcagcagcagtcact 300
DB 1291 agtaagttgcccgcagctgtatcactcgtttagcagcagcagcagtcact 1232
QY 301 gtcacatccatccgtaagatgtcttctgtgactgtgtaactcaaccaatc 360
DB 1231 gtcacatccatccgtaagatgtcttctgtgactgtgtaactcaaccaatc 1172
QY 361 gaatagtgatgcgcgcagcagctgtcttcgtccgcgcgtcaatacggatac 420
DB 1171 gaatagtgatgcgcgcagcagctgtcttcgtccgcgcgtcaatacggatac 1112
QY 421 ccaatagagaagaacttaaaagtctcactcacttgaaaaagttcttgaggaa 480
DB 1111 ccaatagagaagaacttaaaagtctcactcacttgaaaaagttcttgaggaa 1052
QY 481 tcaagatcttaccgctgtgtgagatccagctgcatcaccacgcgcaccacat 540
DB 1051 tcaagatcttaccgctgtgtgagatccagctgcatcaccacgcgcaccacat 992
QY 541 tcttcagcatcttacttaccacagcgttcttggtgagcaaaaacaggaaagc 600
DB 991 tcttcagcatcttacttaccacagcgttcttggtgagcaaaaacaggaaagc 932
QY 601 gcgcgaaaaaggaataagggcgacaggaatgttgaataactactctctctt 660
DB 931 gcgcgaaaaaggaataagggcgacaggaatgttgaataactactctctctt 872
QY 661 caatatctgaagcattatcagggttatgtctcactaagcagatataatgtat 720
DB 871 caatatctgaagcattatcagggttatgtctcactaagcagatataatgtat 812
QY 721 attagaataataaacaataaggggttcgcgcacattcccgaaaaagtcacat 780
DB 811 attagaataataaacaataaggggttcgcgcacattcccgaaaaagtcacat 752
QY 781 gttcaagaaccattatctcactgaataactaataaataagcgtatcaagggcc 840
DB 751 gttcaagaaccattatctcactgaataactaataaataagcgtatcaagggcc 692
QY 841 tttcgcctcgcgcgtttcgtgtatgacgggtgaaaaaccttgacacatgacgtcc 900
DB 691 tttcgcctcgcgcgtttcgtgtatgacgggtgaaaaaccttgacacatgacgtcc 632
QY 901 acggtcacagctgtctgtaagcgcgttcgcgcagcagacagccgctcagggcg 960
DB 631 acggtcacagctgtctgtaagcgcgttcgcgcagcagacagccgctcagggcg 572
QY 961 gcgggtgttgccgcgcgtgtcgcgcgcgttcgcttaactatgcgcgc 1001
DB 571 gcgggtgttgccgcgcgtgtcgcgcgcgttcgcttaactatgcgcgc 531

```

RESULT 3
AAV32980/c

ID AAV32980 standard; DNA; 3190 BP.
 AC AAV32980.
 XX
 XX
 XX 17-NOV-1998 (first entry)
 DE Tn7 target plasmid sequence.
 XX
 XX Tn7; target plasmid; transposon: ATP-utilising regulatory protein;
 KW transposable element; DNA sequencing; genetic analysis;
 KM insertional mutagenesis; ss.
 XX
 XX Synthetic.
 XX
 XX W09837205-A1.
 XX
 XX 27-AUG-1998.
 XX
 XX 20-FEB-1998; 98MO-US03353.
 XX
 XX 20-FEB-1997; 97US-0037955.
 XX
 XX (CRAI/) CRAIG N L.
 XX
 XX Craig NL;
 XX
 XX WPI: 1998-467567/40.
 XX
 XX Transposon(s) encoding mutant ATP using proteins for insertion -
 PT which is efficient and random, with reduced site specificity; for
 PT DNA sequencing and altering gene expression
 XX
 XX
 PS Disclosure; Fig 11B; 143pp; English.
 XX
 XX The present sequence represents a Tn7 target plasmid per 183.
 CC Transposon Tn7 encodes an ATP-utilising regulatory protein that
 CC contains a mutation that allows efficient and simple insertion of,
 CC and reduced target site specificity of, a transposable element derived
 CC from the transposon. The ATP-regulatory protein and compositions are used
 CC for the efficient, non-specific and simple insertion of a transposon or
 CC transposable element into a DNA segment. This is useful in DNA
 CC sequencing, for genetic analysis by insertional mutagenesis, or for
 CC alteration of gene expression by insertion of a desired sequence.
 XX
 XX Sequence 3190 BP; 784 A; 808 C; 773 G; 825 T; 0 other;

Query Match 100.0%; Score 1001; DB 19; Length 3190;
 Best Local Similarity 100.0%; Pred. No. 3,7e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacttatacgcctccatccatcacttaattgttcggaagctagaagt 60
DB 2035 cctgaacttatacgcctccatccatcacttaattgttcggaagctagaagt 1976
QY 61 agtcgcagtaataagtttcgcaagctgttcacattgtaagcagtcgtgtc 120
DB 1975 agtcgcagtaataagtttcgcaagctgttcacattgtaagcagtcgtgtc 1916
QY 121 cgtctcgtttgttgaatgcttcacatccgctcccaagcaagcagttaca 180
DB 1915 cgtctcgtttgttgaatgcttcacatccgctcccaagcaagcagttaca 1856
QY 181 tgatcccccatgtgtgcaaaaagcgtttagctccttcggttcctcgtatc 240
DB 1855 tgatcccccatgtgtgcaaaaagcgtttagctccttcggttcctcgtatc 1796
QY 241 agtaagttgcccgcagctgtatcactcgtttagcagcagcagcagtcact 300
DB 1795 agtaagttgcccgcagctgtatcactcgtttagcagcagcagcagtcact 1736
QY 301 gtcacatccatccgtaagatgtcttctgtgactgtgtaactcaaccaatc 360

```

Db 1735 GTCATGCCATCCGTAAGATGCTTTTCTGTAGTGTGAGTACATCAACCAAGTCAATTCTGA 1676
 QY 361 gaataagtgatgcggcgacggagtgctcttgcggcgctgcaataaggatataccggc 420
 Db 1675 GAATAGGTATGCGGACCGAGTGTCTTGGCCGGGCAATACGGGTAATATACCGCG 1616
 QY 421 ccaatagcagaactctaaagtctcatcatcttgagaaacgtctcttgcggcgcaaaatc 480
 Db 1615 CCACATAGCAGAACTTTAAAGTCTCATCATGTGAAAAAGCTTCTTGGGGCCAAAAATC 1556
 QY 481 tcaagatctaccgctgttgagatccagtlcgtatgtaaccacgtgtgacccaactga 540
 Db 1555 TCAAGGATCTTACCGCTGTGAGATCCAGTTGCATGTAAACCCACTGTCACCCAACTGA 1496
 QY 541 tcttcagcatcttacttctcaccaggttctcgtgtgagcaaaacaggaagcgaat 600
 Db 1495 TCTTCACATCTTTACTTTCACACAGCGTTTCGGGGACAAAAACAGGAAGGCAAAAT 1436
 QY 601 gccgcaaaaaaggaataaggcgacacggaatgttgaatactcaactctctctt 660
 Db 1435 GCCGCAAAAAAGGAATAGGGGACACGGAATGTGATATCACTACTCTCTCTTTT 1376
 QY 661 caatatattgaagcattatcagggatattgtctaatgagcgatatactattgaatg 720
 Db 1375 CAATATTATTGAAGCATTTATVCAAGGGTTATGCTCATGACGCGAATACATATTGAAATG 1316
 QY 721 attagaataataacaataagggtgtccgcacatctcccgcaaaagtgcacactgac 780
 Db 1315 ATTTAGAAAAATAACAAATAGGGGTTCCGCGACATTTCCGAAAAAGTGCACCTTGAC 1256
 QY 781 gtctaaagaaccattatcatcatgacattaaactaaataagcgatcaagagcc 840
 Db 1255 GTCTAGAAACCATTTATCATGACATTAACCTATAAAAATAGGGGTATCAAGAGGCC 1196
 QY 841 tttcgtctgcggcttgcgttgatgacggtgaaacacctgcaacactgacgtccggag 900
 Db 1195 TTTGCTGTCGCGGTTTGGGTGATGACGGTGAACCTCTGACACATGACAGCTCCGGAG 1136
 QY 901 acggtcacaactgtctgtaagcagatgcggagcagaacaacccgtcaaggcggtca 960
 Db 1135 AGGTTACACCTTGTCTGTAGAGGGATGCCGGGAGCAGCAACCCCGTACAGGGCGGTCA 1076
 QY 961 gcgggtgttgaggcggtgtgcgggctggttaactatgcgcgc 1001
 Db 1075 GCGGGTGTGGCGGGGTGCGGCTTGAATTAATGATGCGGC 1035
 RESULT 4
 AAV43439
 ID AAV43439 standard; DNA; 3253 BP.
 XX
 AC AAV43439;
 XX
 DT 29-OCT-1998 (first entry)
 XX
 DE DNA sequence of the vector pGEM3-Zf(-).
 KW recA gene; extein; screening; antimicrobial activity; intein;
 KW protein-splicing element; growth protein; antimicrobial agent;
 KW prevention; splicing; identification; conditionally-splicing; ss.
 OS Synthetic.
 XX
 PN US5795731-A.
 XX
 PD 18-AUG-1998.
 XX
 PF 26-AUG-1996; 96US-0702902.
 XX
 PR 26-AUG-1996; 96US-0702902.
 XX
 PA (HEAL-) HEALTH RES INC.
 XX

PI Belfort M:
 XX
 DR WPI; 1998-46664/40.
 XX
 PT Screening agents for antimicrobial activity - by monitoring affect
 PT of splicing intein into reporter gene, also useful for studying
 PT intein function
 PS Example 1; Fig 6B; 64pp; English.
 XX
 CC The present sequence represents the DNA sequence of the vector
 CC pGEM3-Zf(-), used in the course of the invention. The specification
 CC describes a method for screening agents for activity against a microbial
 CC pathogen that has an intein (i.e. a protein-splicing element occurring
 CC naturally as an in-frame protein fusion) in a gene that encodes a protein
 CC that facilitates growth. The method comprises preparing recombinant
 CC clones of an inducible expression vector containing an altered reporter
 CC gene, including a silent restriction site, and the intein, and detecting
 CC the presence of the extein product of the intein by recombinant clones in
 CC presence of varying concentrations of the agent. Reduced production of
 CC the extein product indicates inhibition of the intein, i.e. that the
 CC agent has antimicrobial activity. The agents of the invention represent
 CC a new type of antimicrobial agent that prevents splicing out of the
 CC intein and thus formation of an active, essential protein. The same
 CC method is used to identify conditionally-splicing inteins and to study
 CC intein function.
 SQ Sequence 3253 BP; 793 A; 817 C; 805 G; 838 T; 0 other:

Query Match 100.0%; Score 1001; DB 19; Length 3253;
 Best Local Similarity 100.0%; Pred. No. 3, 7e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttataccgctccatccacgtatattgttgcgggaagctagaagt 60
 Db 1533 cctgaacttataccgctccatccacgtatattgttgcgggaagctagaagt 1592
 QY 61 agtgcgaagtaataagttgcgaaggtgttgccattgttaaggatcgtgtgtca 120
 Db 1593 agtgcgaagtaataagttgcgaaggtgttgccattgttaaggatcgtgtgtca 1652
 QY 121 cgtctgcttctgtatgtcttcaatcagctccggttcccaacgataagcgagttaca 180
 Db 1653 cgtctgcttctgtatgtcttcaatcagctccggttcccaacgataagcgagttaca 1712
 QY 181 tgaatcccatgtgtgtgcaaaaaagcggttagctcccttggtctccgactgtgtcaga 240
 Db 1713 tgaatcccatgtgtgtgcaaaaaagcggttagctcccttggtctccgactgtgtcaga 1772
 QY 241 agtaagttgcgcgaagtttatcatcatgtgtatgtgcaagactgacataattcttact 300
 Db 1773 agtaagttgcgcgaagtttatcatcatgtgtatgtgcaagactgacataattcttact 1832
 QY 301 gtcatgcatacgttaagatgtcttctgtgactgtgtgacttaacaaagtcattctga 360
 Db 1833 gtcatgcatacgttaagatgtcttctgtgactgtgtgacttaacaaagtcattctga 1892
 QY 361 gaataagttatagcggaacagagttgtcttcccggtcgaataagggataaccggc 420
 Db 1893 gaataagttatagcggaacagagttgtcttcccggtcgaataagggataaccggc 1952
 QY 421 ccacatagcagaactttaaagtgtcatatattggaagaagttcttcggggcgaaatc 480
 Db 1953 ccacatagcagaactttaaagtgtcatatattggaagaagttcttcggggcgaaatc 2012
 QY 481 tcaagatcttaccgctgttgagatccagttcgatgtaaccactcgttgacccaactga 540
 Db 2013 tcaagatcttaccgctgttgagatccagttcgatgtaaccactcgttgacccaactga 2072
 QY 541 tcttcagcatcttacttccaccagcgttctcgtgtgagcaaaaaacaggaagcgaat 600
 Db 2073 tcttcagcatcttacttccaccagcgttctcgtgtgagcaaaaaacaggaagcgaat 2132

QY 601 gccgcaaaaaggaataagggcgacacggaagaatgttgaatactactcttctt 660
 |||||
 Db 2133 gccgcaaaaaggaataagggcgacacggaagaatgttgaatactactcttctt 2192
 QY 661 caatattattgaagcattatcgaaggtattgtctcctatgagcgatcatattgaatg 720
 |||||
 Db 2193 caatattattgaagcattatcgaaggtattgtctcctatgagcgatcatattgaatg 2252
 QY 721 attagaataaatacaataaggggttcgcgcacattcccccgaagaatgtccactgac 780
 |||||
 Db 2253 attagaataaatacaataaggggttcgcgcacattcccccgaagaatgtccactgac 2312
 QY 781 gtctaaagaacattatattatcatgacatttaacctataaataaggggtatccagaagccc 840
 |||||
 Db 2313 gtctaaagaacattatattatcatgacatttaacctataaataaggggtatccagaagccc 2372
 QY 841 ttctgtctgcgcgttcctgcgtgacacggtgaaacctctgacacatgacgtctcccgag 900
 |||||
 Db 2373 ttctgtctgcgcgttcctgcgtgacacggtgaaacctctgacacatgacgtctcccgag 2432
 QY 901 acggtcacagctgtctgttaagcggatcgcgagacagacaaagccgtcagggcgctca 960
 |||||
 Db 2433 acggtcacagctgtctgttaagcggatcgcgagacagacaaagccgtcagggcgctca 2492
 QY 961 gcgaggtttggcggtgtcgcggggtcgtgcttaactatgagc 1001
 |||||
 Db 2493 gcgaggtttggcggtgtcgcggggtcgtgcttaactatgagc 2533

RESULT 5

AA299245/C

ID AA299245 standard; DNA: 3331 BP.

AC AA299245;

DT 03-JUL-2000 (first entry)

DE Nucleotide sequence of a Pinpoint expression vector.

KW Pinpoint vector; fusion protein antigen; membrane antigen; syphilis; ss.

XX OS Treponema pallidum.

XX PN EP985931-A2.

XX PD 15-MAR-2000.

XX PF 12-AUG-1999; 99EP-0115877.

XX PR 04-SEP-1998; 98US-0148920.

XX PA (BECTON DICKINSON & CO.

XX PI Mullenix MC, Deutsch J;

XX DR WPI: 2000-226057/20.

PT Immunoassay for anti-Treponema pallidum antibodies, used for diagnosis
 of syphilis using a fusion protein of membrane antigen with peptide
 sequence that can be biotinylated in vivo -

PS Disclosure: Page 8-9; 16pp: English.

XX The present sequence represents the Pinpoint vector, which is used in
 CC the course of the invention. The specification describes a method for
 CC detecting antibodies against Treponema pallidum. The antibodies are
 CC detected in a sample by reaction with a fusion protein antigen, present
 CC in the mixture in limiting concentration. The fusion protein antigen
 CC comprises a Treponema pallidum membrane antigen. The method is used
 CC for diagnosis of syphilis.

SQ Sequence 3331 BP; 811 A; 854 C; 885 G; 781 T; 0 other;

Query Match	100.0%	Score 1001	DB 21	Length 3331
Best Local Similarity	100.0%	Pred. No. 3	7e-263	
Matches 1001	Conservative	0	Mismatches 0	Indels 0
Gaps 0				
QY 1 cctgcaacttataccgcctccatccagtcattatattgttcgaggaagctagagtaag 60				
Db 1565 cctgcaacttataccgcctccatccagtcattatattgttcgaggaagctagagtaag 1506				
QY 61 agtgcagtaataagatttgcgcaagctgtgtccatgtctaacagcgatcgtgtgtca 120				
Db 1505 agtgcagtaataagatttgcgcaagctgtgtccatgtctaacagcgatcgtgtgtca 1446				
QY 121 cgtctgt 180				
Db 1445 cgtctgt 1386				
QY 181 tgatcccccattgtgtgcaaaaaagcggttaagctccctgcctccatcgtgtgtca 240				
Db 1385 tgatcccccattgtgtgcaaaaaagcggttaagctccctgcctccatcgtgtgtca 1326				
QY 241 agtaagtgtgcgagtgatctatcatcgaatgtgtatgtgcagacatgataatcttact 300				
Db 1325 agtaagtgtgcgagtgatctatcatcgaatgtgtatgtgcagacatgataatcttact 1266				
QY 301 gttatccatccgttaagatgcttctgtgactgtgtgtgtgtgtgtgtgtgtgtgtgt 360				
Db 1265 gttatccatccgttaagatgcttctgtgactgtgtgtgtgtgtgtgtgtgtgtgtgt 1206				
QY 361 gaatagttatacgcgcgcagctgtctctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420				
Db 1205 gaatagttatacgcgcgcagctgtctctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1146				
QY 421 ccaatagcagaactttaaagtgtcacaatcattgtgaaacggttcttcgagggagaaac 480				
Db 1145 ccaatagcagaactttaaagtgtcacaatcattgtgaaacggttcttcgagggagaaac 1086				
QY 481 tcaagatcttaccgctgt 540				
Db 1085 tcaagatcttaccgctgt 1026				
QY 541 tcttcagcatcttacttacttaccagcggttctgtgtgtgtgtgtgtgtgtgtgtgtgt 600				
Db 1025 tcttcagcatcttacttacttaccagcggttctgtgtgtgtgtgtgtgtgtgtgtgtgt 966				
QY 601 gccgcaaaaaggaataagggcgacacggaagaatgttgaatactactcttctt 660				
Db 965 gccgcaaaaaggaataagggcgacacggaagaatgttgaatactactcttctt 906				
QY 661 caatattattgaagcattatcgaaggtattgtctcctatgagcgatcatattgaatg 720				
Db 905 caatattattgaagcattatcgaaggtattgtctcctatgagcgatcatattgaatg 846				
QY 721 attagaataaatacaataaggggttcgcgcacattcccccgaagaatgtccactgac 780				
Db 845 attagaataaatacaataaggggttcgcgcacattcccccgaagaatgtccactgac 786				
QY 781 gtctaaagaacattatattatcatgacatttaacctataaataaggggtatccagaagccc 840				
Db 785 gtctaaagaacattatattatcatgacatttaacctataaataaggggtatccagaagccc 726				
QY 841 ttctgtctgcgcgttcctgcgtgacacggtgaaacctctgacacatgacgtctcccgag 900				
Db 725 ttctgtctgcgcgttcctgcgtgacacggtgaaacctctgacacatgacgtctcccgag 666				
QY 901 acggtcacagctgtctgttaagcggatcgcgagacagacaaagccgtcagggcgctca 960				
Db 665 acggtcacagctgtctgttaagcggatcgcgagacagacaaagccgtcagggcgctca 606				
QY 961 gcgaggtttggcggtgtcgcggggtcgtgcttaactatgagc 1001				
Db 605 gcgaggtttggcggtgtcgcggggtcgtgcttaactatgagc 565				

Query Match	100.0%	Score 1001	DB 11	Length 3343
Best Local Similarity	100.0%	Pred. No. 3.8e-263		
Matches 1001	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	cclygcacattlalcgcgcctccalcacagttacatataatgtlbgccggaagcagtagaagtaagt	60		
Db 2188	CCGTCACACTTATATCCGCCCTCCATCCAGTCTATTATTTGTTGCCGGGAAGCTAGAGTAAGT	2129		
QY 61	agttgcgcagtaataatggtttgcgaacggtttgtgcatttgttaagaagcaccgfyggttca	120		
Db 2128	AGTTGCGCAATTATATGTTTGGCGCAACGTTGTTGCCATTGTCTAAAGCATCGTGTGTTCA	2069		
QY 121	cagtcgtcgttttgatagtgtcttcattcagctccgcgtttcccaagatcaagcgagttaca	180		
Db 2068	CGCTCGTCGTTTGATGTATGCTTCATTCAAGCTCCGGTGTCCAAAGCATCAAGCGAGTTACA	2009		
QY 181	tgtatcccccatgtttgycgcaaaaacggtttgagctcccttggtccctccagatcgtttgtcaga	240		
Db 2008	TGATATCCCATGTTGTGCAAAAAACGGTGTATGCTTCCTTCCTCCGATCGTTGTACAG	1949		
QY 241	agaagaattggccgcagtgltatcacctcaagtgtatgacgacgcgtcaatactctactact	300		
Db 1948	AGTAAGTTGGCGCAAGTATCATCATCATGTTATGAGCAGACACTGCATTAATTTCTTACT	1889		

QY	301	gtcaagcatccgtaagaatgcttctcgtgtaacgtggaagtactcaacaaatgcatcttga	360
Db	1888	gtcaagcatccgtaagaatgcttctcgtgtaacgtggaagtactcaacaaatgcatcttga	1829
QY	361	gaatagtgtaatgcgagccagagttgcctcttgcgcgcgtcaatacggataataccgcg	420
Db	1828	gaatagttatgcggcgacacggagctgcttgcgcgcgtcaatacggagatattaccgcg	1769
QY	421	cccaatgacgaactttaaaagtgcttcacatcatttggaaaaagcttcttcgggcgaaatc	480
Db	1768	ccacattacacaaactttttaaagtgcttcacatcatttggaaaaagcttcttcgggcgaaatc	1709
QY	481	tcaaggaattcaacgcgttcttgaatccagttgagtttaacccacgcgtgcacccaactga	540
Db	1708	tcaaggaattcaacgcgttcttgaatccagttgagtttaacccacgcgtgcacccaactga	1649
QY	541	lcttcagcatcttacttcttccaccagcgttcttcggtgtagcaaaaaacagaaagcnaat	600
Db	1648	tcttcagcatcttacttcttccaccagcgttcttcggtgtagcaaaaaacagaaagcnaat	1589
QY	601	gcccgaataaaaggaataaagcgcgacaaggaataatgttgaatactatactcttcccttt	660
Db	1588	gcccgaataaaaggaataaagcgcgacaaggaataatgttgaatactatactcttcccttt	1529
QY	661	caatattatgaagcatcttcaagggttatttcttcataagagggatataatttgaatgt	720
Db	1528	caatattatgaagcatcttcaagggttatttcttcataagagggatataatttgaatgt	1469
QY	721	attagaataaataacaataaagaggttcgcgcgcacatttcccccgaagaagtgcacactgac	780
Db	1468	attagaataaataacaataaagaggttcgcgcgcacatttcccccgaagaagtgcacactgac	1409
QY	781	gtcttaagaagaacattatattcatatgacattaaacctataaataaggcgtatcaagaagcc	840
Db	1408	gtcttaagaagaacattatattcatatgacattaaacctataaataaggcgtatcaagaagcc	1349
QY	841	tttcgtctcgcgcgttcttcggtgaatgcaggttgaacaccccttgcacataagctccggag	900
Db	1348	tttcgtctcgcgcgttcttcggtgaatgcaggttgaacaccccttgcacataagctccggag	1289
QY	901	acggtcacacgcttgcctcgttaagaacgagaatgcgcggaagcaagaacccgcgtcaaggcgctca	960
Db	1288	acggtcacacgcttgcctcgttaagaacgagaatgcgcggaagcaagaacccgcgtcaaggcgctca	1229
QY	961	gcggggttgcggcggttgcgtgcgggctgcggttgaactatgcgcg	1001
Db	1228	gcggggttgcggcggttgcgtgcgggctgcggttgaactatgcgcg	1188
RESULT	7		
AAAS1632/c			
ID	AAAS1632	standard; DNA; 3383 BP.	
XX	AAAS1632:		
AC	AAAS1632:		
XX			
DT	31-Oct-2000	(first entry)	
DE			
XX			
XX			
KW	Uridine phosphorylase: udp: purine nucleoside phosphorylase; PNP: deod:		
KW	Genetically modified organism; catalytic; transglycosylation;		
KW	nucleoside analogue; alpha-pentose-1-phosphate sugar; phosphorolysis;		
KW	anti-viral; anti-tumour; mesophilic bacterium; ss.		
XX			
OS	Escherichia coli.		
OS	Synthetic.		
XX			
FH	Key	Location/Qualifiers	
FT	misc_RNA	1..230	
FT		/tag= a	
FT		/label= pUC18_sequence	
FT	CDS	216..952	
FT		/tag= b	
TT			

Ft		/product= lacZ-deod_fused_gene
Ft	misc_RNA	961..3383
Pt		/+tag- C
Tt		/label= puc18_sequence
Xx		
Pn		
Pd	MW200039307-A2.	
Xx		
Pd	06-JUL-2000.	
Xx		
Pf	23-DEC-1999;	99MO-EPI0416.
Xx		
Pr	23-DEC-1998;	98IT-MI02792.
Xx	(NORP-) NORPHARMA SPA.	
Pa		
Xx	Betetti G, Call' S, Ghisotti D, Orsini G, Tonon G, Zuffi G,	
Ei		
Dl	WPJ; 2000-452402/39.	
Dr		
Xx		
Pt	Recombinant expression vectors used to transform cells for the production of e.g. nucleosides encode uridine phosphorylase and/or purine nucleoside phosphorylase	
Pf		
Xx	Claim 14; Page 47-48; 72pp; English.	
Ps		
Xx		
Cc	Host cells genetically modified to express uridine phosphorylase (udp)	
Cc	and purine nucleoside phosphorylase (PNP) or their corresponding crude or purified extracts, either separately or in combination are used as catalysts of transglycosylation reactions between a donor nucleoside and an acceptor base, for preparing nucleoside analogues containing heterocyclic systems with purine and/or pyrimidine bases substituted by one or more nitrogen atoms, for preparing alpha-pentose-1-phosphate sugars by phosphorolysis reactions and for producing nucleosides and modified nucleoside analogues (all claimed). The modified or natural CC nucleosides are used directly or as intermediates in the preparation of drugs with anti-viral and anti-tumour activity and for preparing oligonucleotides for therapeutic or diagnostic use. The recombinant host CC cells are produced using recombinant plasmid expression vectors comprising at least one gene sequence of a mesophilic bacterium coding for a polypeptide with udp activity and/or enzyme PNP activity and at least one upe gene sequence coding for antibiotic resistance. In particular, CC the E. coli uid and deo genes are used. Recombinant strains produced CC using the vectors express poly-peptides with enzyme UDP and PNP activity CC in large amounts, e.g. 340-1040 times higher UDP activity and 120-200 CC times higher PNP activity than non-transformed corresponding wild type CC strains.	
Sq	Sequence 3383 BP; 826 A; 868 C; 855 G; 834 T; 0 other;	
	Query Match	100.0%; Score 1001; DB 21; Length 3383;
	Best Local Similarity	100.0%; Pred. No. 3,8e-263;
	Matches 1001; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Yy	I ccgcgaacttacccgcccctcacatcaagtcitaaattgtttcgsgagaagttagaagtaagr 60 	
Db	2228 CCTGCAACTTATACCGCCCTTCATTCCAGTGCTATTATTTGTGCCGGGAAGCTAAGTAAGT 2169 	
Oy	61 agtcgccagtaaataaggtttgacacaagtlbtggscalcgtgtcaaaggaacatgcyytgtcca 120 	
Db	2168 AGTTCCGCACAATAATAGTTTGGCAGAACGTMGTGGCATTTGCTAACAGGCATCGHNGHTCCA 2109 	
Oy	121 cgcctgtgcgtttgatatgatgtcatcataagtcaccggtltccccaaagataaaggcgagttacca 180 	
Db	2108 CGCGTCGTCGTTTTGTAATGAGCTTCATTCAGCTCCGGTTCACCAGCATPAAGCGAAGTTTACA 2049 	
Oy	181 tgaatccccaagtgtgtgcacaaaagaacgggtlaagcctcccttcggluccctcgacacctgtgtcaga 240 	
Db	2048 TGATCCCCCATGTTGTGTGCAAAAAAACGGGTGTAGCTOCTTCGGSTCTCCGATCGTGTGCAGA 1989 	
Oy	241 actgaagtgtgcccagatgttatcactaatggtatagcagaacatgtaalctctctaact 300 	
Db	1988 AGTAAGTAGGGCCGCAATGTTATCACCTCATGGTATATGAGACAGCATCAAATAATCTCTTACT 1929 	

Oy	301	gtctgcacacccgaagaagctcttcgtgactggttgactcaaacagctcatctga	360
Db	1928	GTCTGCCATCCGTAAGATGCTTTTCTTGAGACTGGTGAATCTCAACCAATCTTTTGA	1869
Oy	361	gaatagtgatctgcggcgaccgaagctgtctcttgcgccgcgctcaatacggataaataccgcg	420
Db	1868	GAATAGTGTATGGCGGACGACGAGTGTCTCTTGCCCGGGGTCAATACGGGATTAATACGGCG	1809
Oy	421	ccacatagcagaactttaaagtgtctcatcalttgtaaaaagcttctcgggcgaaaact	480
Db	1808	CCACATGTGCAAGAACTTTAAAAGTGCTCATCATTTGGAAAAAGCTTCTCGGGGCGAAAACTC	1749
Oy	481	tcaagatcttaccgcggttggaatccagtttgatgtgaacccactcgtgaccccaatga	540
Db	1748	TCAAGATCTTACCGGCTGTTGAGATCAAGTGTGAATGTAAACCACTCGGACCCAACTGA	1689
Oy	541	tcttcagcatcttctaacttccaccagcgcttctgggtgtagcaaaaacagaaagcaaat	600
Db	1688	TCTTCAACATCTTTTACTTTCACACACGTTCTTGGGTGAAGCAAAACGAGAGCAAAAT	1629
Oy	601	gcgcgcaaaaaaggaataaaggcgacacggaatgtgtgaatactcaactcttcctltt	660
Db	1628	GCCGCCAAAAAGGGAATTAAGGGGCGACACGGAATGTTAAATACATCAATCTTCTTTT	1569
Oy	661	caatatattgaagcagcttatacaggttatttctccatgaagcgatataattgaatgt	720
Db	1568	CAATATTATTGAAGCATTTATTCAGGGTTTATTGTCATGAGGAGATCAATTTTGAATGT	1509
Oy	721	attagaaaaaataaacaatagagggtccgcgacatctcccgaaaagtccacctgac	780
Db	1508	ATTTAGAAAAATTAACAATAATAGGGGTTCCGGCCACATTTCCCGCAAAAGTGCACCTGAC	1449
Oy	781	gtctaaagaaccattatcatgacataaccataaataagcglatcaagagccc	840
Db	1448	GTCMAAAMAAACCATTTATTATCATGACATTTAAACCTTAAATAATAGCGATCACAGGCCCC	1389
Oy	841	ttctgctcgcggttctcggtatgaacggttgaaaacccctgaacatgagctcccgag	900
Db	1388	TTTCTGCTCCGGCGTTTCCGTGATGACGGGTGAAAACCTGTACACATGACAGCTCCGGAG	1329
Oy	901	acggtcacagcttctgtctgtaacgcgatagcgcggagcagacaagaaccgctcagggcgctca	960
Db	1328	ACGGTACACACTTGTCTGTAAAGGGGATGCGCGGACACAGAACGCCCGTCAAGGGGGGTCA	1269
Oy	961	gcggggtgtgcggggtgtctgcgggcttgcttaactatgcgc	1001
Db	1268	GCGGGTGTTGGCGGGTGTCGGGGCTGACTTAACATATGCGGC	1228
RESULT 8			
AAAS1630/c			
ID	AAAS1630 standard; DNA; 3444 BP.		
XX	XX		
AC	AAAS1630;		
XX	XX		
DE	31-OCT-2000 (first entry)		
XX	XX		
DE	Plasmid pGM679 containing E. coli udp gene.		
XX	XX		
KW	uridine phosphorylase; udp; purine nucleoside phosphorylase; PNP; deod;		
KW	Genetically modified organism; catalyzt; transglycosylation;		
KW	nucleoside analogue; alpha-pentose-1-phosphate sugar; phosphorolysis;		
KW	anti-viral; anti-tumour; mesophilic bacterium; ss.		
XX	XX		
OS	Escherichia coli.		
XX	XX		
OS	Synthetic.		
XX	XX		
FT	Key	Location/Qualifiers	
FT	misc_RNA	1..242	
FT		/tag= a	
FT		/label= PDC18_sequence	
FT	CDS	216..1013	

FT		/+tag= b
FT		/product= lacZ-udp_fused_gene
FT		1022..3444
FT	misc_RNA	/+tag= c
FT		/label= pUC18_sequence
PX	WO200039307-A2.	
XX		
PD	06-JUL-2000.	
PX		
PF	23-DEC-1999;	99WO-EPI0416.
XX		
PR	23-DEC-1998;	98IT-MI02792.
PX	(NORP-) NORPHARMA SPA.	
PA		
PI	Bestetti G, Calli' S, Ghisotti D, Orsini G, Tonon G, Zuffi G;	
XX		
DR	WPI: 2000-452402/39.	
XX		
PT	Recombinant expression vectors used to transform cells for the	
PT	production of e.g. nucleosides encode uridine phosphorylase and/or	
PT	purine nucleoside phosphorylase	
XX		
PS	Claim 14; page 43-44; 72pp; English.	
XX		
CC	Host cells genetically modified to express uridine phosphorylase (udp)	
CC	and purine nucleoside phosphorylase (PNP) or their corresponding crude or	
CC	purified extracts, either separately or in combination are used as	
CC	catalysts of transglycosylation reactions between a donor nucleoside and	
CC	an acceptor base, for preparing nucleoside analogues containing	
CC	heterocyclic systems with purine and/or pyrimidine bases substituted by	
CC	one or more nitrogen atoms, for preparing alpha-pentose-1-phosphate	
CC	sugars by phosphorolysis reactions and for producing nucleosides and	
CC	modified nucleoside analogues (all claimed). The modified or natural	
CC	nucleosides are used directly or as intermediates in the preparation of	
CC	drugs with anti-viral and anti-tumour activity and for preparing	
CC	oligonucleotides for therapeutic or diagnostic use. The recombinant host	
CC	cells are produced using recombinant plasmid expression vectors	
CC	comprising at least one gene sequence of a mesophilic bacterium coding	
CC	for a polypeptide with udp activity and/or enzyme PNP activity and at	
CC	least one gene sequence coding for antibiotic resistance. In particular,	
CC	the E. coli udp and deoD genes are used. Recombinant strains produced	
CC	using the vectors express polypeptides with enzyme Udp and PNP activity	
CC	in large amounts, e.g. 340-1040 times higher udp activity and 120-200	
CC	times higher PNP activity than non-transformed corresponding wild type	
CC	strains.	
XX		
SQ	Sequence 3444 BP; 820 A; 881 C; 887 G; 856 T; 0 other:	

Query Match	100.0%	Score 1001	DB 21	Length 3444
Best Local Similarity	100.0%	Pred. No. 3.8e-263		
Matches 1001	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0
QY	1	cctgcacattatccgcgcccatccagttcatataattgttcgcgggaagctagagtaagt	60	
Db	2289	cctgcacattattttccgccttcccatccagtcagtatttaattgtttgcgggaagctagagtaagt	2230	
QY	61	agttgcacagtaataatagtttgcgcaagcttgttgcattgtctacaggaatcgttgtgtca	120	
Db	2229	agttgcacagtaataatagtttgcgcaagcttgttgcattgtctacaggaatcgttgtgtca	2170	
QY	121	gcctgtcgttctgtgatacttcattcaacgtccgcgttcccaacgaatcaaggcgaagtaca	180	
Db	2169	gcctgtcgttctgtgatacttcattcaacgtccgcgttcccaacgaatcaaggcgaagtaca	2110	
QY	181	tgatcccccatttctgtgcaaaaaagcgtgtagctcccttgcgttccttcgactcgtttgcaga	240	
Db	2109	tgatcccccatttctgtgcaaaaaagcgtgtagctcccttgcgttccttcgactcgtttgcaga	2050	
QY	241	agtaaatgttgcgcagtgatataactatcattgttataatgtgcagcaactgcataattcttaact	300	

Db	2049	AGTAAGTTGGCCGACAGTGTATACATCAGTGTTATGCGACAGCTGCATTAATCTCTTACT	1990
OY	301	gtcatgcataccgtaagaatgctttcttgtaactggtgtaactcaaccaatgatacttcga	360
Db	1989	GTCAATGCCATCCGTAAGATCATCTTTCTGTGTACGTGGAGTACTTCAACCAAGCATTTCTGA	19300
OY	361	gaataatgatacggcggaaccgaatgtctcttcgcggcggtcaatacggataataaccgcg	420
Db	1929	GAATAGGTGTATGGCGGCACCGAGTTGCTCTTGCCCGCGGTCAATAGGATTAATACCGCG	18707
OY	421	ccacatagcagaactcttaaaagtctcatcatatgaaaaagctcttcgcggcgcaaaatc	480
Db	1869	CCACATTAACCAAACTTTAAAGTGTCAATCATTTGAAAAAGTTCTTCCGGCGCAAACTTC	18110
OY	481	tcaaggatctaacgcgtgttgagataccagttcgaatgtaaaccaactcgtgaccccaatcga	540
Db	1809	TCAAGATCTTACCGCTGTGAGATCCAGTTCCGATTAACCCACTCTGTCCACCCAACTGA	17500
OY	541	tcttcagatcttttactcttcacacaggtttcgggtlgaacaaaaacgaagaagcaaaat	600
Db	1749	TCTTCAGCATCTTTTACTTTTCACACAGGTTCTGGGTGAGCAAAAACAGAAAGGCAAAAT	16900
OY	601	gcgcgaaaaaagggaataaggcgacacggaatgtctgaatactcaactcttccttlt	660
Db	1689	GCCGCAAAAAGGGAATAAGGGCGACACGSAATGTTGAACTACTACTCTTCCTTTTT	16300
OY	661	caatatattgaagcatttalcagaggtlaltgtctcatgagcggatacataltgaatgt	720
Db	1629	CAATATTTTAAACATTTTATCAGGGTTATTGTCTCATAGCGGATACATATTTGAATGT	15707
OY	721	attagaataataacaataaggggttcgcgcgaactttcccggaagaatgtccactctgac	780
Db	1569	ATTTAGAAAAATAAAACAATAAGGGGTTCCGCGACATTTTCCCGAAAATGTCCACTGTAC	15100
OY	781	gtctaaagaacattatcatcatatgacataaacataaagaatagcgtlctcaagcgcc	840
Db	1509	GCTTAGAAACATTATTTATCATGACATTAACTATTAAAAATAAGCGTATCAGAGGCC	14500
OY	841	tttcgtctcgcggttttcogtgaatgacggtgaaacaccttgacaacatcagctcccgag	900
Db	1449	TTTGCTGTCCGGCGTTTCGGTGATGAGACGGTGAAAAACCTCTGCACATCATGAGCTCCCGAG	13900
OY	901	acggtcaacagcttgtctgttaagcgtatgcccgggaagcagacaagaacccgttaaggcggtca	960
Db	1389	AAGGTCACACAGCTTCTGTGTAAAGCGAGACCGGGAGAGACAAACCCCTGTAGGGCCGCTCA	13300
OY	961	gcgggtgtctgcgggtgttcggagtcgtgacttaacatagcgcc	1001
Db	1329	GGGGGTGTGGCGGGGTGTGGGGGTGGCTTAACCTATATCGCGC	1289

RESULT	9
AAAT78825	
ID	AAAT78825 standard; DNA; 3819 BP.
XX	
AC	AAAT78825;
XX	
DT	23-JAN-1998 (first entry)
XX	
DE	Kappa light chain plasmid pIc6C5.
XX	
KW	Ig; affinity constant; human; antigen; hybridoma; B cell; transgene;
KW	transgenic; mouse; CD4; antibody; autoimmune; inflammatory;
KW	transplant rejection; immunoglobulin; ss.
XX	
OS	Synthetic.
OS	Homo sapiens.
XX	
PN	W09713852-A1.
XX	
PD	17-APR-1997.
XX	
PF	10-OCT-1996; 96WO-US16433.

XX 10-OCT-1995; 9505-0544404.
 PR (GENP-) GENPHARM INT. INC.
 XX
 XX Kay RM, Lonberg N;
 PI
 XX WPI: 1997-235888/21.
 DR
 XX Novel anti-CD4 antibody produced by transgenic mice - used in the
 PT treatment of auto-immune disease etc.
 XX
 PS Example 42; Page 266-268; 396pp; English.
 XX
 CC A novel composition has been developed which comprises an immunoglobulin
 CC (Ig) having an affinity constant (K_a) of at least 2 multiply
 CC 100000000 M⁻¹ for binding to a predetermined human antigen. The
 CC present sequence represents the kappa light chain plasmid pLC6G5.
 CC which includes the kappa constant region and polyadenylation site. Anti-
 CC CD4 antibodies may be used in therapeutic and diagnostic applications,
 CC especially for the treatment of human diseases. These antibodies reduce
 CC activity of CD4 cells and reduce undesirable autoimmune reactions.
 CC inflammatory response and transplant rejection. Transgenic animals are
 CC capable of producing heterologous antibodies of multiple isotypes by
 CC undergoing isotype switching. These animals produce a first Ig type
 CC that is necessary for antigen-stimulated B-cell maturation and can
 CC switch to encode and produce one or more subsequent heterologous
 CC isotypes.
 CC
 CC Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other;
 SO

Query Match 100.0%; Score 1001; DB 18; Length 3819;
 Best Local Similarity 100.0%; Pred. No. 3.9e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaactatccgcctccatccatcagctcttaattgttcggaagctagaagtaagt 60
 DB 1156 cctgaactatccgcctccatccatcagctcttaattgttcggaagctagaagtaagt 1215
 QY 61 agtgcagcttaagttgtgcgaacgtgtgtccattgtctaaagcattcgtgtgtca 120
 DB 1216 agtgcagcttaagttgtgcgaacgtgtgtccattgtctaaagcattcgtgtgtca 1275
 QY 121 cgtcgtcgtgtgtgtatgtgtctatcagctccggttcccaagatcaagcgagttaca 180
 DB 1276 cgtcgtcgtgtgtgtatgtgtctatcagctccggttcccaagatcaagcgagttaca 1335
 QY 181 tgatccccaatgtgtgcgaacacggttagctccttcggtccgcgtcgtgtgtca 240
 DB 1336 tgatccccaatgtgtgcgaacacggttagctccttcggtccgcgtcgtgtgtca 1395
 QY 241 agtaagttgcgcagctgttactactatgtgtatgtgcagcagctgcataatcttact 300
 DB 1396 agtaagttgcgcagctgttactactatgtgtatgtgcagcagctgcataatcttact 1455
 QY 301 gtcatgcatccgttaagatgtcttctgtgactgtgtgtactcaaccaatcattctga 360
 DB 1456 gtcatgcatccgttaagatgtcttctgtgactgtgtgtactcaaccaatcattctga 1515
 QY 361 gaatagtgatcgcgcgcagctgtgtccttcgcccgcggtcaatacggataacgcg 420
 DB 1516 gaatagtgatcgcgcgcagctgtgtccttcgcccgcggtcaatacggataacgcg 1575
 QY 421 ccaataagaagaacttaaaagtgtctacatctgtgaaaacgttctcggcggaanaact 480
 DB 1576 ccaataagaagaacttaaaagtgtctacatctgtgaaaacgttctcggcggaanaact 1635
 QY 481 tcaagaattaccgctgtgtgagatccagtttgatgaataccacactcgtgcacccaactga 540
 DB 1636 tcaagaattaccgctgtgtgagatccagtttgatgaataccacactcgtgcacccaactga 1695
 QY 541 tcttcagcatcttacttaccacagcgttctcgtgtgagcaaaaacagggaagcaaat 600

DB 1696 tcttcagcatcttacttaccacagcgttctcgtgtgagcaaaaacagggaagcaaat 1755
 QY 601 gccgcaaaaaagggaataaggcgacacaggaatgttgaatactactcttctttt 660
 DB 1756 gccgcaaaaaagggaataaggcgacacaggaatgttgaatactactcttctttt 1815
 QY 661 caatatattgaagcatctatacagggttatgttcataagcagatattgaattgt 720
 DB 1816 caatatattgaagcatctatacagggttatgttcataagcagatattgaattgt 1875
 QY 721 attagaataataacaataaggggttcgcgcacattcccggaanaagtgccactgac 780
 DB 1876 attagaataataacaataaggggttcgcgcacattcccggaanaagtgccactgac 1935
 QY 781 gtctaagaacacattatcatatgacatcaactataaataagtgatcagggccc 840
 DB 1936 gtctaagaacacattatcatatgacatcaactataaataagtgatcagggccc 1995
 QY 841 ttctgcctcgcggttcgtgtgtagcgtgtgaaaacctctgcacacatgcagctccggag 900
 DB 1996 ttctgcctcgcggttcgtgtgtagcgtgtgaaaacctctgcacacatgcagctccggag 2055
 QY 901 acggtcacagctgtgtgtgaagcgatgcgcggagcagacaaagccggtcaaggcgctga 960
 DB 2056 acggtcacagctgtgtgtgaagcgatgcgcggagcagacaaagccggtcaaggcgctga 2115
 QY 961 gcgggtgtgtgcgggtgtgtgcgggctgtgttaactatgcgc 1001
 DB 2116 gcgggtgtgtgcgggtgtgtgcgggctgtgttaactatgcgc 2156

RESULT 10

AAV39266
 ID AAV39266 standard; DNA; 3819 BP.

XX AAV39266;
 AC 18-DEC-1998 (first entry)
 DT
 XX Plasmid pLC6G5 nucleotide sequence.
 DE
 XX Transgenic animal; human heterologous antibody; transgene;
 KW isotype switching; neutrophil efflux; reperfusion injury; CD4 binding;
 KW autoimmune reaction; inflammatory response; transplant rejection;
 KW acid induced lung injury; acute adult respiratory distress syndrome;
 KW AIDS; vasculitis; septic shock; allergic reaction; asthma;
 KW cystic fibrosis; ss.
 XX
 XX Synthetic.
 OS Homo sapiens.
 OS
 XX WO9824884-A1.
 PN 11-JUN-1998.
 PD
 XX 01-DEC-1997; 97MO-US21803.
 PF
 XX 02-DEC-1996; 96US-0758417.
 PR
 XX (GENP-) GENPHARM INT.
 PA
 XX Kay RM, Lonberg N;
 PI
 XX WPI: 1998-333306/29.
 DR
 XX Hybridoma producing antibody specific for interleukin-8 - used to
 PT prevent efflux of neutrophils from vasculature, and treat
 PT reperfusion injury
 PS Example 42; Pages 317-319; 452pp; English.
 XX The present sequence represents a plasmid, pLC6G5, which contains a

CC synthetic kappa light chain sequence (created using oligonucleotide
CC AAV39244-65). This synthetic sequence differs from natural sequences in
CC that strings of repeated oligonucleotides are interrupted (to facilitate
CC oligonucleotide synthesis and PCR amplification), optimal translation
CC initiation sites are incorporated and HindII sites were engineered
CC upstream of the translation initiation sites. The plasmid is used in the
CC construction of minigenes for expression of Igkappa anti-CD4
CC antibodies, in the transgenic mouse of the invention. The specification
CC describes transgenic non-human animals, especially a mouse, which are
CC capable of producing a human heterologous antibodies of multiple isotypes
CC by undergoing isotype switching. The transgenic animals have human heavy
CC and light chain transgenes. The transgenes are capable of functionally
CC rearranging a heterologous diversity (D) gene in a
CC variable-diversity-junction (V-D-J) recombination. The transgenes
CC include a heavy chain transgene comprising at least one V, D and J gene
CC segment, and one constant region gene segment. The immunoglobulin (Ig)
CC light chain transgene comprises at least one V and J gene segment and one
CC constant region gene segment. The gene segments are heterologous to the
CC transgenic animal. The antibody can be used to prevent efflux of
CC neutrophils from vasculature. It can also be used to treat reperfusion
CC injury. CD4 binding antibodies are used to reduce undesirable autoimmune
CC reactions, inflammatory responses and rejection of transplanted organs.
CC The anti-IL-8 antibodies can reduce tissue damage and prolong survival
CC in animal models of acute adult respiratory distress syndrome (ARDS) and
CC acid induced lung injury. The anti-IL-8 antibodies can also be used for
CC the treatment of vasculitis, septic shock, allergic reactions
CC (e.g. asthma) and cystic fibrosis.

XX Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other;

Query Match 100.0%; Score 1001; DB 19; Length 3819;
Best Local Similarity 100.0%; Pred. No. 3.9e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaactttatcgctccatcagcttattatgttgcgggaagctagaaagt 60
DB 1156 cctgaactttatcgctccatcagcttattatgttgcgggaagctagaaagt 1215
QY 61 agtgcgcagttatagtttgcaagctgtgtccatgtgcaagagatctgtgtca 120
DB 1216 agtgcgcagttatagtttgcaagctgtgtccatgtgcaagagatctgtgtca 1275
QY 121 cgtcgtcgttctgtatagcttcatcagctcgggtcccaagataagcgaaattca 180
DB 1276 cgtcgtcgttctgtatagcttcatcagctcgggtcccaagataagcgaaattca 1335
QY 181 tgatcccatgttctgtcaaaaagcggttagctccctcgtcctcgatcgtgtcaga 240
DB 1336 tgatcccatgttctgtcaaaaagcggttagctccctcgtcctcgatcgtgtcaga 1395
QY 241 agtaagttgcgcaggtgtatcatcattgttatgcaagcactgataattccttact 300
DB 1396 agtaagttgcgcaggtgtatcatcattgttatgcaagcactgataattccttact 1455
QY 301 gtcattgcacatccgaatagatcttctgtgactgtgtgactcaacaagtcattctga 360
DB 1456 gtcattgcacatccgaatagatcttctgtgactgtgtgactcaacaagtcattctga 1515
QY 361 gaatagttatgcgcgcagaggtgtgtcttgcgcgcggttaatacgggataatccgcg 420
DB 1516 gaatagttatgcgcgcagaggtgtgtcttgcgcgcggttaatacgggataatccgcg 1575
QY 421 ccacatacagaactttaaaagtgtcatcattgtgaaaagcttctcggggcgaaactc 480
DB 1576 ccacatacagaactttaaaagtgtcatcattgtgaaaagcttctcggggcgaaactc 1635
QY 481 tcaagatcttaccgctgttgagatccagttcgatgttaaccactcgtgcacccaactga 540
DB 1636 tcaagatcttaccgctgttgagatccagttcgatgttaaccactcgtgcacccaactga 1695
QY 541 tcttaagatcttacttaccacgaggttctcgtgtgagcaaaaacaggaagcaaat 600
DB 1696 tcttaagatcttacttaccacgaggttctcgtgtgagcaaaaacaggaagcaaat 1755

DB 1696 tcttaagatcttacttaccacgaggttctcgtgtgagcaaaaacaggaagcaaat 1755
QY 601 gccgcaaaaagggaataaggcgacacggaatgttgaatactactctctctttt 660
DB 1756 gccgcaaaaagggaataaggcgacacggaatgttgaatactactctctctttt 1815
QY 661 caattattgaagattatcaggttatgtctcgaagcggatcatattgaatgt 720
DB 1816 caattattgaagattatcaggttatgtctcgaagcggatcatattgaatgt 1875
QY 721 attagaanaataacaatagaggttcgcgcacattcccgaaaagtgccaccgac 780
DB 1876 attagaanaataacaatagaggttcgcgcacattcccgaaaagtgccaccgac 1935
QY 781 gttcaagaacattatattatcagacattcaactataaataagcggtatcagagccc 840
DB 1936 gttcaagaacattatattatcagacattcaactataaataagcggtatcagagccc 1995
QY 841 ttctgcgcgcgtttcgtgtagcaggtgaanaaccttgacacatcgactccggag 900
DB 1996 ttctgcgcgcgtttcgtgtagcaggtgaanaaccttgacacatcgactccggag 2055
QY 901 acggtacagcttctgttgaagcggatgcggagcagacacagccgttcaggcgctca 960
DB 2056 acggtacagcttctgttgaagcggatgcggagcagacacagccgttcaggcgctca 2115
QY 961 gccggtgttgccgggtgtcggggtcgtgcttaactacatcgagc 1001
DB 2116 gccggtgttgccgggtgtcggggtcgtgcttaactacatcgagc 2156

RESULT 11

AAZ22020
ID AAZ22020 standard; DNA: 3819 BP.

AAZ22020:

24-NOV-1999 (first entry)

XX Nucleotide sequence of plasmid pUC65.

XX Transgenic animal; heterologous antibody; hybridoma; B cell;

XX Transgenic mouse; human heavy chain transgene; digoxin;

XX human light chain transgene; immortalized cell; immunoglobulin;

XX Shiga-like toxin; autoimmune disease; cancer; infectious disease;

XX transplant rejection; blood disorder; coagulation disorder; ss.

XX Synthetic.

XX WO9445962-A1.

XX 16-SEP-1999.

XX 12-MAR-1999; 99WO-US05535.

XX 13-MAR-1998; 98US-0042353.

XX (GENP-) GENPHARM INT. INC.

XX Lonberg N, Fishwild DM, Ball WJ;

XX WPI; 1999-551219/46.

XX Novel transgenic non-human animals used to produce heterologous

XX antibodies

XX Example 42; Page 318-320; 484pp; English.

XX The specification describes transgenic animals that are capable of
XX producing a heterologous antibody. The antibodies are isolated from a
XX hybridoma, comprising B cells, that is obtained from a transgenic mouse
XX having a genome comprising a human heavy chain transgene and a human
XX light chain transgene. The B cells are fused to immortalized cells

CC suitable for generating a hybridoma, which produces a detectable
CC amount of an immunoglobulin that specifically binds digoxin or
CC Shiga-like toxin. B cells from transgenic animals can be used to
CC generate hybridomas expressing monoclonal high affinity human sequence
CC antibodies. Antibodies produced from the transgenic animals of the
CC invention can be used to treat human diseases, e.g. autoimmune
CC diseases, cancer, infectious disease, transplant rejection, blood
CC disorders such as coagulation disorders and other diseases. The
CC present sequence is used in the course of the invention.

SQ Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other;

Query Match: 100.0%; Score 1001; DB 20; Length 3819;	
Best Local Similarity 100.0%; Pred. No. 3, 9e-263;	
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 cctgcgaacttaccgcctccaccagctcattaatgtgttcgcgggaagctagaagta 60
Db	1156 cctgcgaacttaccgcctccaccagctcattaatgtgttcgcgggaagctagaagta 1215
QY	61 agttcgcaattaatagtttgcgcgaacggtgtgtgcacattgcttaaggatcgtgttca 120
Db	1216 agttcgcaattaatagtttgcgcgaacggtgtgtgtgcacattgcttaaggatcgtgttca 1275
QY	121 cgcctgcgcgtttgtatgcttcattcacaactccgcgtttcccaacgatcaaggcgattaca 180
Db	1276 cgcctgcgcgtttgtatgcttcattcacaactccgcgtttcccaacgatcaaggcgattaca 1335
QY	181 tgaatccccatgtttgtgcaaaaaaagcggttagcctccctcgcgttcctcgatccgttgttata 240
Db	1336 tgaatccccatgtttgtgcaaaaaaagcggttagcctccctcgcgttcctcgatccgttgttata 1395
QY	241 agtaagtcggccgcgaagtgtatatacccaacggtgtatggaacacgtgcataacttccttact 300
Db	1396 agtaagtcggccgcgaagtgtatatacccaacggtgtatggaacacgtgcataacttccttact 1455
QY	301 gtcatcgccatccgtaagaatgctttctcgtgtactggttgsagtactcaaccaagtcatcttga 360
Db	1456 gtcatcgccatccgtaagaatgctttctcgtgtactggttgsagtactcaaccaagtcatcttga 1515
QY	361 gaataagtgatctggcgggcgaacccgaattgcctttgccggcgctcaataccggataataaccgc 420
Db	1516 gaataagtgatctggcgggcgaacccgaattgcctttgccggcgctcaataccggataataaccgc 1575
QY	421 ccacatagcgaagaactttaaaagtgcccatcatttggaaaacgctttcttcggggcgaaaaatc 480
Db	1576 ccacatagcgaagaactttaaaagtgcccatcatttggaaaacgctttcttcggggcgaaaaatc 1635
QY	481 tcaaggaatcttaacgcgctgttgagataccagtttcgatgttaaccacatcggtgcacccaactga 540
Db	1636 tcaaggaatcttaacgcgctgttgagataccagtttcgatgttaaccacatcggtgcacccaactga 1695
QY	541 tcttcagacatcttattactttcacccagcggtttccgggttgagcaaaaaacggaaagcgaatat 600
Db	1696 tcttcagacatcttattactttcacccagcggtttccgggttgagcaaaaaacggaaagcgaatat 1755
QY	601 gccgcgaaaaaaagggaataaaggcgcacacggaatgtttgataatcctaactcttcctttt 660
Db	1756 gccgcgaaaaaaagggaataaaggcgcacacggaatgtttgataatcctaactcttcctttt 1815
QY	661 caatatattgaacacatttaccaggtttatgttccatataaggcggataaattttgaagt 720
Db	1816 caatatattgaacacatttaccaggtttatgttccatataaggcggataaattttgaagt 1875
QY	721 atttagaaaaataaacaataaaggggttccgcgcgacatttcgccgaaaaagtgcacacttgac 780
Db	1876 atttagaaaaataaacaataaaggggttccgcgcgacatttcgccgaaaaagtgcacacttgac 1935
QY	781 gttctaaagaacacattatcatcatgacataaactataaaaaataggcgtatatacgaagccc 840
Db	1936 gttctaaagaacacattatcatcatgacataaactataaaaaataggcgtatatacgaagccc 1995

[illegible]

RESULT 12

ID	standard; DNA; 3858 BP.
AAD04947	

AC AAD04947;

DT 17-JUL-2001 (first entry)

DE Plasmid pRK50 used to test Cre recombinase mediated inversion.

KW Gene trapping construct; conditional mutation; unidirectional inversion;

KW selection cassette; transgenic organism; pRK50; Cre recombinase

OS Chimeric - Adenovirus.

OS chimeric - ECMV virus.

PN WO200129208-A1.

PD 26-APR-2001.

PF 16-OCT-2000; 2000WO-EP10162.

PR	16-OCT-1999;	99EP-0120592.
PR	27-OCT-1999.	99TS-0162016.

XX
XX
(APPE-) APPEMIS PHARM GMBH
DA

PA (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
XX

PI Kuehn R, Von Melchner H, Altschmied J,
vv

DR WPI; 2001-308486/32.

PT New gene trapping construct capable of causing conditional mutations in
PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped -
XX
PS Example 3; Page 73-74; 78pp: English.

Example 3; Page 73-74; 78pp; English.

The present invention relates to a conditional gene trapping construct capable of causing conditional mutations in genes. The gene trapping construct comprises two functional DNA segments, each being flanked by two recombinase recognition sequences (RRS) specific to site specific recombinase which is capable of unidirectional inversion of double standard DNA segment. One of the DNA segment (disruption cassette) is inserted in antisense orientation relative to the transcriptional orientation of the gene to be trapped. The other DNA segment (selection cassette) is inserted in sense direction relative to the transcriptional orientation of the gene to be trapped. The cell comprising the gene trapping construct is useful for the identification and/or isolation of genes. The transgenic organism comprising the gene trapping construct is useful to study gene function at various developmental stages. The gene trapping construct is useful for mutationally inactivating all cellular genes. The present sequence is PK50 vector, which is used to test Cre recombinase mediated inversion, which is related to the invention.

Sequence 3858 BP; 943 A; 959 C; 994 G; 962 T; 0 other;


```

Db 1156 cctgcacattatccgcctccatccagctctattatattgttcgcgggaagctagaagtc 1215
Qy 61 agtcgcacgttaataagtttgcgcacagctgttcgcacattgtctcaaggacgcgtgtgta 120
Db 1216 agtcgcacgttaataagtttgcgcacagctgttcgcacattgtctcaaggacgcgtgtgta 1275
Qy 121 cgcctgcctgttgcgttaagcttcatcagctccggttcccaagatcaaggcaggttaca 180
Db 1276 cgcctgcctgttgcgttaagcttcatcagctccggttcccaagatcaaggcaggttaca 1335
Qy 181 tgatcccccatgttgcgttaaaaaaagcggttagctccctcgtccctcgcacgttgcaca 240
Db 1336 tgatcccccatgttgcgttaaaaaaagcggttagctccctcgtccctcgcacgttgcaca 1395
Qy 241 agtaagttgcgcgcaggttatactcaatcagttatagcagacactgtcaataattctctact 300
Db 1396 agtaagttgcgcgcaggttatactcaatcagttatagcagacactgtcaataattctctact 1455
Qy 301 gtcatagcacatcgttaagaatgctttctgtgacgtgtgagtaactcaaccagatctctga 360
Db 1456 gtcatagcacatcgttaagaatgctttctgtgacgtgtgagtaactcaaccagatctctga 1515
Qy 361 gaataagttatgcgcgcagcaggttgccttcgtccgcgcgtcaatacgggataataccgcg 420
Db 1516 gaataagttatgcgcgcagcaggttgccttcgtccgcgcgtcaatacgggataataccgcg 1575
Qy 421 ccacatagcagaactttaaaagttgcatactatgtaaaacgttcttcggggcgaaacatc 480
Db 1576 ccacatagcagaactttaaaagttgcatactatgtaaaacgttcttcggggcgaaacatc 1635
Qy 481 tcaagatcttaacgcgtgttgatgacagttcgatgttaacccactcgtgcacccaatgta 540
Db 1636 tcaagatcttaacgcgtgttgatgacagttcgatgttaacccactcgtgcacccaatgta 1695
Qy 541 tcttcacagatcttacttaccacagcgttctcgttggtgagcaaaaaaagcaaggcaaat 600
Db 1696 tcttcacagatcttacttaccacagcgttctcgttggtgagcaaaaaaagcaaggcaaat 1755
Qy 601 ggcgcgaaaaaaggaataagggcgacacgcgaacatgttgaatactatactctctttt 660
Db 1756 ggcgcgaaaaaaggaataagggcgacacgcgaacatgttgaatactatactctctttt 1815
Qy 661 caataatgtgaagacattatcagaggttatgtctcatgagcggtatatactattggaatg 720
Db 1816 caataatgtgaagacattatcagaggttatgtctcatgagcggtatatactattggaatg 1875
Qy 721 attagaaaaataaacaataaggggttcgcgcacattcccccgaagaagtccacctgac 780
Db 1876 attagaaaaataaacaataaggggttcgcgcacattcccccgaagaagtccacctgac 1935
Qy 781 gctcaagaacaacattatatactatgacattaaactataaaataagcggtatcagagggcc 840
Db 1936 gctcaagaacaacattatatactatgacattaaactataaaataagcggtatcagagggcc 1995
Qy 841 ttctgctcgcgcgttcttcgtgtgagcagcggtgtaaaacctctgaacatgcaactcccgag 900
Db 1996 ttctgctcgcgcgttcttcgtgtgagcagcggtgtaaaacctctgaacatgcaactcccgag 2055
Qy 901 acggtcacacagctgtcttaagcgagatccgggagcaagaacccgtcaaggcggtca 960
Db 2056 acggtcacacagctgtcttaagcgagatccgggagcaagaacccgtcaaggcggtca 2115
Qy 961 gcgggtgttgcgcgggtgtcgcggggtcgtgcttaactatgcgcg 1001
Db 2116 gcgggtgttgcgcgggtgtcgcggggtcgtgcttaactatgcgcg 2156

```

```

RESULT 14
AA78801
ID AA78801 standard; DNA; 3881 BP.
XX
AC AA78801;
XX

```

```

DT 23-JAN-1998 (first entry)
XX
DE Kappa light chain plasmid pCK7-96.
XX
KW Ig; affinity constant; human; antigen; hybridoma; B cell; transgene;
KW transgenic; mouse; CD4; antibody; autoimmunity; inflammatory;
KW transplant rejection; immunoglobulin; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN MO9713852-A1.
XX
PD 17-APR-1997.
XX
PF 10-OCT-1996; 96WO-US16433.
XX
PR 10-OCT-1995; 95US-0544404.
XX
PA (GENP-) GENPHARM INT INC.
XX
PI Kay RM, Lonberg N;
XX
PS WPI; 1997-235888/21.
XX
PT Novel anti-CD4 antibody produced by transgenic mice - used in the
PT treatment of auto-immune disease etc.
XX
Example 42; Page 260-262; 396pp; English.
XX
CC A novel composition has been developed which comprises an immunoglobulin
CC (Ig) having an affinity constant (Ka) of at least 2 multiply
CC 1000000000 M-1 for binding to a predetermined human antigen 96
CC which includes the kappa constant region and polyadenylation site. Anti-
CC CD4 antibodies may be used in therapeutic and diagnostic applications,
CC especially for the treatment of human diseases. These antibodies reduce
CC activity of CD4 cells and reduce undesirable autoimmune reactions,
CC inflammatory response and transplant rejection. Transgenic animals are
CC capable of producing heterologous antibodies of multiple isotypes by
CC undergoing isotype switching. These animals produce a first Ig type
CC that is necessary for antigen-stimulated B-cell maturation and can
CC switch to encode and produce one or more subsequent heterologous
CC isotypes.
XX
SQ Sequence 3881 BP; 981 A; 1014 C; 921 G; 965 T; 0 other:

```

```

Query Match 100.0%; Score 1001; DB 18; Length 3881;
Best Local Similarity 100.0%; Pred. No. 4e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 cctgcacattatccgcctccatccagctctattatattgttcgcgggaagctagaagtc 60
Db 1156 cctgcacattatccgcctccatccagctctattatattgttcgcgggaagctagaagtc 1215
Qy 61 agtcgcacgttaataagtttgcgcacagctgttcgcacattgtctcaaggacgcgtgtgta 120
Db 1216 agtcgcacgttaataagtttgcgcacagctgttcgcacattgtctcaaggacgcgtgtgta 1275
Qy 121 cgcctgcctgttgcgttaagcttcatcagctccggttcccaagatcaaggcaggttaca 180
Db 1276 cgcctgcctgttgcgttaagcttcatcagctccggttcccaagatcaaggcaggttaca 1335
Qy 181 tgatcccccatgttgcgttaaaaaaagcggttagctccctcgtccctcgcacgttgcaca 240
Db 1336 tgatcccccatgttgcgttaaaaaaagcggttagctccctcgtccctcgcacgttgcaca 1395
Qy 241 agtaagttgcgcgcaggttatactcaatcagttatagcagacactgtcaataattctctact 300
Db 1396 agtaagttgcgcgcaggttatactcaatcagttatagcagacactgtcaataattctctact 1455
Qy 301 gtcatagcacatcgttgaagaatgctttctgtgacgtgtgagtaactcaaccagatctctga 360

```

```

DB 1456 gtcacgcatcgcgaatgcttctctgacgctgagctacccacgaatctctga 1515
QY 361 gaatagtgtgtgctgacgacgagcttctctgcccggcgtcaatagcgataaccgcg 420
DB 1516 gaatagtgtgtgctgacgacgagcttctctgcccggcgtcaatagcgataaccgcg 1575
QY 421 ccaatagcagaacttaaaagctcaltcatctgaaacgcttctccggcgcaaaacc 480
DB 1576 ccaatagcagaacttaaaagctcaltcatctgaaacgcttctccggcgcaaaacc 1635
QY 481 tcaagatcttaccgctgttgaatccagttcgaatgaacccactgtgacccaactga 540
DB 1636 tcaagatcttaccgctgttgaatccagttcgaatgaacccactgtgacccaactga 1695
QY 541 tcttcagcatcttcttaccgagcttctggttgagcgcaaaacaggaagcgcaaat 600
DB 1696 tcttcagcatcttcttaccgagcttctggttgagcgcaaaacaggaagcgcaaat 1755
QY 601 gccgcaaaaaggaataaggcgacaggaatgttgaatactactactctcttctt 660
DB 1756 gccgcaaaaaggaataaggcgacaggaatgttgaatactactactctcttctt 1815
QY 661 caatattatgaagcattatcaagggttattgtctctgaagcgacatacttgaatg 720
DB 1816 caatattatgaagcattatcaagggttattgtctctgaagcgacatacttgaatg 1875
QY 721 attagaataataacaataagaggttcgacgacattccccgaaatgtccactgac 780
DB 1876 attagaataataacaataagaggttcgacgacattccccgaaatgtccactgac 1935
QY 781 gtctaaagaacattatcatctgaatataactataaaatagcgcttcacagagccc 840
DB 1936 gtctaaagaacattatcatctgaatataactataaaatagcgcttcacagagccc 1995
QY 841 ttctgcctcgcgcttctggtatgacggttgaaaccccttgacacatcagctcccgag 900
DB 1996 ttctgcctcgcgcttctggtatgacggttgaaaccccttgacacatcagctcccgag 2055
QY 901 acggtacagctgtctgtgaagcggaatgcgcgagacagacagccggtacagcgctga 960
DB 2056 acggtacagctgtctgtgaagcggaatgcgcgagacagacagccggtacagcgctga 2115
QY 961 gcgggtgtgtgcgggtgtcgcgggcttgcttaactatgcgcgc 1001
DB 2116 gcgggtgtgtgcgggtgtcgcgggcttgcttaactatgcgcgc 2156

RESULT 15
AAZ21996
ID AAZ21996 standard; DNA; 3881 BP.
XX
AC AAZ21996;
XX
DT 24-NOV-1999 (first entry)
DE
XX Nucleotide sequence of plasmid pCK7-'96.
DE
XX Transgenic animal; heterologous antibody; hybridoma; B cell;
KW transgenic mouse; human heavy chain transgene; digoxin;
KW human light chain transgene; immortalized cell; immunoglobulin;
KW Shiga-like toxin; autoimmune disease; cancer; infectious disease;
KW transplant rejection; blood disorder; coagulation disorder; ss.
XX
OS Synthetic.
XX
XX WO9945962-A1.
XX
XX 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US05535.
XX
XX 13-MAR-1998; 98US-0042353.
PR

```

```

XX
PA (GENP-) GENPHARM INT INC.
XX
PI Lonberg N, Fishwild DM, Ball WJ;
XX
DR WPI: 1999-551219/46.
XX
PT Novel transgenic non-human animals used to produce heterologous
PT antibodies
XX
PS Example 42; Page 311-313; 484pp; English.
XX
CC The specification describes transgenic animals that are capable of
CC producing a heterologous antibody. The antibodies are isolated from a
CC hybridoma, comprising B cells, that is obtained from a transgenic mouse
CC having a genome comprising a human heavy chain transgene and a human
CC light chain transgene. The B cells are fused to immortalized cells
CC suitable for generating a hybridoma, which produces a detectable
CC amount of an immunoglobulin that specifically binds digoxin or
CC Shiga-like toxin. B cells from transgenic animals can be used to
CC generate hybridomas expressing monoclonal high affinity human sequence
CC antibodies. Antibodies produced from the transgenic animals of the
CC invention can be used to treat human diseases, e.g. autoimmune
CC diseases, cancer, infectious disease, transplant rejection, blood
CC disorders such as coagulation disorders and other diseases. The
CC present sequence is used in the course of the invention.
XX
SQ Sequence 3881 BP; 981 A; 1014 C; 921 G; 965 T; 0 other;

```

```

Query Match 100.0%; Score 1001; DB 20; Length 3881;
Best Local Similarity 100.0%; Pred. No. 4e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 cctgaacttattccgctccatccagctctatattgttgcgggaagctagaatg 60
DB 1156 cctgaacttattccgctccatccagctctatattgttgcgggaagctagaatg 1215
QY 61 agtcgcagcttaagtgttcgcaagctgttgccttctgaagcagctcgtgtga 120
DB 1216 agtcgcagcttaagtgttcgcaagctgttgccttctgaagcagctcgtgtga 1275
QY 121 cgtctcgtctgttgaatgcttcaatcagctccggtcccaacagatcaagcgagtaca 180
DB 1276 cgtctcgtctgttgaatgcttcaatcagctccggtcccaacagatcaagcgagtaca 1335
QY 181 tgatcccccagctgtgtcgaaaaagcggttagctcctctggtccctccgatcgtgtcaga 240
DB 1336 tgatcccccagctgtgtcgaaaaagcggttagctcctctggtccctccgatcgtgtcaga 1395
QY 241 agtaagtgtgcgcagctgttatctcatatgttatgacacatcgtatcttctact 300
DB 1396 agtaagtgtgcgcagctgttatctcatatgttatgacacatcgtatcttctact 1455
QY 301 gtcatgcacatcgtlaaagatcttctgtgacgtgtgagtaactcaacaagatcatctga 360
DB 1456 gtcatgcacatcgtlaaagatcttctgtgacgtgtgagtaactcaacaagatcatctga 1515
QY 361 gaatagtgtatgcggacgagcttcttgcggcggtcaatacgggataaccgcg 420
DB 1516 gaatagtgtatgcggacgagcttcttgcggcggtcaatacgggataaccgcg 1575
QY 421 ccaatagcagaacttaaaagctcaltcatctgaaacgcttctccggcgcaaaacc 480
DB 1576 ccaatagcagaacttaaaagctcaltcatctgaaacgcttctccggcgcaaaacc 1635
QY 481 tcaagatcttaccgctgttgaatccagttcgaatgaacccactgtgacccaactga 540
DB 1636 tcaagatcttaccgctgttgaatccagttcgaatgaacccactgtgacccaactga 1695
QY 541 tcttcagcatcttcttaccgagcttctggttgagcgcaaaacaggaagcgcaaat 600
DB 1696 tcttcagcatcttcttaccgagcttctggttgagcgcaaaacaggaagcgcaaat 1755

```

QY 601 gccgcaaaaaagggaataaggcgacacggaatgttgaatactactactcttctt 660
|||||
Db 1756 gccgcaaaaaagggaataaggcgacacggaatgttgaatactactactcttctt 1815
|||||
QY 661 caataatttgaagcattatcagaaggttatgtctcatgagcgatatacatatttgaatgt 720
|||||
Db 1816 caataatttgaagcattatcagaaggttatgtctcatgagcgatatacatatttgaatgt 1875
|||||
QY 721 attagaataaaataaagaatagggttccgcgcacatttcccgaaaaagtgcacactgac 780
|||||
Db 1876 attagaataaaataaagaatagggttccgcgcacatttcccgaaaaagtgcacactgac 1935
|||||
QY 781 gcttaagaataccattatcatatgacattaaactataaaataaggcgtatcatcagaagccc 840
|||||
Db 1936 gcttaagaataccattatcatatgacattaaactataaaataaggcgtatcatcagaagccc 1995
|||||
QY 841 ttctgctcgcgcgttttcggtgatgacggtgaaaaactctgacacatgcagctcccgag 900
|||||
Db 1996 ttctgctcgcgcgttttcggtgatgacggtgaaaaactctgacacatgcagctcccgag 2055
|||||
QY 901 acggtcacagcttctctgtaagcgatcccggaagcagacacagcccgtaagggcgctca 960
|||||
Db 2056 acggtcacagcttctctgtaagcgatcccggaagcagacacagcccgtaagggcgctca 2115
|||||
QY 961 gcggggtgttgcggggtgtcggggctgcttaactatgcgcgc 1001
|||||
Db 2116 gcggggtgttgcggggtgtcggggctgcttaactatgcgcgc 2156
|||||

Search completed: January 17, 2002, 16:33:16
Job time: 12501 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:26:40 ; Search time 118.22 Seconds
(without alignments)
1917.649 Million cell updates/sec

Title: US-09-740-211-13_COPY_6000_7000

Perfect score: 1001
1 cctgcacttaccgcctc.....gctgccttactatgcgcgc 1001

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 segs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	3104	1 US-07-415-307A-1	Sequence 1, App1
2	1001	100.0	3104	1 US-08-371-320-1	Sequence 1, App1
3	1001	100.0	3343	6 5453363-2	Patent No. 5453363
4	1001	100.0	3819	4 US-09-042-353-393	Sequence 393, App
5	1001	100.0	3819	4 US-08-758-417A-243	Sequence 243, App
6	1001	100.0	3875	4 US-09-039-982A-21	Sequence 21, App1
7	1001	100.0	3875	4 US-09-039-641-21	Sequence 21, App1
8	1001	100.0	3875	4 US-09-039-762A-21	Sequence 21, App1
9	1001	100.0	3878	4 US-09-039-982A-27	Sequence 27, App1
10	1001	100.0	3878	4 US-09-039-641-27	Sequence 27, App1
11	1001	100.0	3878	4 US-09-039-762A-27	Sequence 27, App1
12	1001	100.0	3881	4 US-09-042-353-369	Sequence 369, App
13	1001	100.0	3881	4 US-08-758-417A-217	Sequence 217, App
14	1001	100.0	3883	4 US-09-039-982A-30	Sequence 30, App1
15	1001	100.0	3883	4 US-09-039-641-30	Sequence 30, App1
16	1001	100.0	3883	4 US-09-039-762A-30	Sequence 30, App1
17	1001	100.0	3908	4 US-09-039-982A-24	Sequence 24, App1
18	1001	100.0	3908	4 US-09-039-641-24	Sequence 24, App1
19	1001	100.0	3908	4 US-09-039-762A-24	Sequence 24, App1
20	1001	100.0	4045	4 US-08-464-700-54	Sequence 54, App1
21	1001	100.0	4273	3 US-08-795-430-3	Sequence 3, App1
22	1001	100.0	4622	4 US-08-509-024-6	Sequence 6, App1
23	1001	100.0	4622	4 US-09-333-279-6	Sequence 6, App1
24	1001	100.0	4696	2 US-08-929-967-15	Sequence 15, App1
25	1001	100.0	4723	4 US-09-042-353-370	Sequence 370, App
26	1001	100.0	4723	4 US-08-758-417A-218	Sequence 218, App
27	1001	100.0	4926	4 US-09-042-353-418	Sequence 418, App

28	1001	100.0	4926	4 US-08-758-417A-268	Sequence 268, App
29	1001	100.0	4950	2 US-08-789-333F-58	Sequence 58, App1
30	1001	100.0	5158	2 US-08-929-967-16	Sequence 16, App1
31	1001	100.0	5181	3 US-08-801-344-5	Sequence 5, App1
32	1001	100.0	5181	4 US-09-498-599-5	Sequence 5, App1
33	1001	100.0	5292	2 US-08-793-610-3	Sequence 3, App1
34	1001	100.0	5294	2 US-08-793-610-2	Sequence 2, App1
35	1001	100.0	5323	2 US-08-793-610-1	Sequence 1, App1
36	1001	100.0	5364	2 US-08-793-610-4	Sequence 4, App1
37	1001	100.0	5639	3 US-09-175-690A-1	Sequence 1, App1
38	1001	100.0	5865	4 US-09-011-745-8	Sequence 8, App1
39	1001	100.0	6028	4 US-09-011-745-5	Sequence 5, App1
40	1001	100.0	6061	4 US-09-011-745-6	Sequence 6, App1
41	1001	100.0	6253	3 US-08-893-327-15	Sequence 15, App1
42	1001	100.0	6280	3 US-08-893-327-17	Sequence 17, App1
43	1001	100.0	6280	3 US-08-893-327-19	Sequence 19, App1
44	1001	100.0	6295	2 US-08-659-206A-4	Sequence 4, App1
45	1001	100.0	6306	5 PCT-US94-00658-1	Sequence 1, App1

ALIGNMENTS

RESULT 1
US-07-415-307A-1
Sequence 1, Application US/07415307A
Patent No. 5344757
GENERAL INFORMATION:
APPLICANT: Holte, Hans Joachim
APPLICANT: Seidl, Rudolf
APPLICANT: Schmitz, Gudrun
APPLICANT: Scholer, Hans
APPLICANT: Kessler, Christoph
APPLICANT: Matthes, Ralf
TITLE OF INVENTION: Process for the Detection of Nucleic Acids
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC/DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/415,307A
FILING DATE: 1990109
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP89/0026
FILING DATE: 12-Jan-1989
APPLICATION NUMBER: DE 38 00 642.1
FILING DATE: 12-Jan-1988
APPLICATION NUMBER: DE 38 13 278.8
FILING DATE: 20-Apr-1988
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5344757man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3104 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-415-307A-1

Query Match 100.0%; Score 1001; DB 1; Length 3104;
 Best Local Similarity 100.0%; Pred. No. 2.1e-301;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttlatccgcctccatccagctcattatgtgtccgggaagctagaagtaagt 60
 |||||||
 Db 1549 CCTGCACACTTATCCGCGCTCATCCAGTATTAATTGTTCCGGGAGAGCTACAGTAAGT 1608
 |||||||

QY 61 agtgcgcagtaataagttgacgaacgttgttgcattgtctacaggaatcgttgtca 120
 |||||||
 Db 1609 AGTTCGCCAGTTAATAGTTGGCCCAACGTTGTGCAATTCCTACAGGCACTCGGTGCTCA 1668
 |||||||

QY 121 cgcctcgcgtttgttattggttattcagctcaggttcccaagatcaaggcgagttaca 180
 |||||||
 Db 1669 CGCTCGTCGTTGTGTAAGGCTTATTCAGCTCCGGTTCCTCCAGATCAAGGCGAGTTACA 1728
 |||||||

QY 181 tgatcccccatgtgtgcaaaaaaagcgtttagctcctcgtctccgactcgttgtcaga 240
 |||||||
 Db 1729 TGATCCCCCATGTTGTGCAAAAAAGGGGTTAGCTCCTTCGCTCCTCGATCGTTGCAGA 1788
 |||||||

QY 241 agtaagttgacgacgttgtatactcaatggttattgacgacactgataattccttact 300
 |||||||
 Db 1789 AGTAAGTTGGCCGACGTGTATCATCATGCTTATGCGACACTGCATTAATTCCTTACT 1848
 |||||||

QY 301 gtcacgcaatcgcgttaagatgcttcttctgactggtgagtaactcaaccgaatctctga 360
 |||||||
 Db 1849 GTCATGCGATCCGTAAGATGCTTTCTTGACTGGTGAAGTACTCAACGCAAGTATTCTGA 1908
 |||||||

QY 361 gaatagtatagcgcgcacgacaggttgtctctgcccgcgtlcaatacgggaataccgcg 420
 |||||||
 Db 1909 GAATAGTGTATGCGGACGACGAGTCTCTTGCGCGCGTCATACGGAATATACCGCG 1968
 |||||||

QY 421 ccaactatagaacacttaaaagcgtcattgtgaaaaacgttctcggcggaacacc 480
 |||||||
 Db 1969 CCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAACGTTCTTCGGGGGGAACATC 2028
 |||||||

QY 481 tcaaggatcttaacgctgttgagatccagtctgagttaaccacactcgttaccaccaactga 540
 |||||||
 Db 2029 TCAAGATGTTACCGCTGTTGAGATGCACTTCGATGTAAACCATGCTGACCCCACTGA 2088
 |||||||

QY 541 tcttcagcatcttacttcttaccacagcttctcgtgtgagcaaaaaacaggaaagcaaat 600
 |||||||
 Db 2089 TCTTCAGCATCTTTACTTTCAACACGCTTCTCGGTGAGCAAAAAACAGAGGCAAAAT 2148
 |||||||

QY 601 gccgcgaaaaaagggaataaggcgacacgaaatgttgaatactatactcttcttctt 660
 |||||||
 Db 2149 GCCGCAAAAAAGGAAATAGGGGACACGGAATGTGAATACTCATACTCTTCTTTT 2208
 |||||||

QY 661 caatatattgaagcatatatcaggttatgttctcatagcggaatactatttgaatgt 720
 |||||||
 Db 2209 CAATATTATTGAAGCATTTATCAGGGTTATGTCTCATGAGCGGATACATATTTAATGT 2268
 |||||||

QY 721 atttagaaaaaataaacaataggggttcgcgcacacattccccgaaaaagtgcacactgac 780
 |||||||
 Db 2269 ATTATCAAAAAATAACAATAAGGGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGAC 2328
 |||||||

QY 781 gttctaaagaacattatttcatgaacttaaaccttaaaaataggcgtatatacagggccc 840
 |||||||
 Db 2329 GTCTAAGAAACATTTATTTATCATGACATTAACCTTAATAAAATAGCGGTATCTCGAGGCC 2388
 |||||||

QY 841 ttctcgtctcgcgcttctcgttgtatagcgttgaaaaacctctgacacatgcagctccggag 900
 |||||||
 Db 2389 TTTTCGTTCTGCGCGCTTTGCGTGATGACGGGTGAACACCTCTGACATCATGAGTCCCGGAG 2448
 |||||||

QY 901 aacgtgtcaacgcttctcgttaagcggatgcgggagcagaacagcccgctcagggcgagctca 960
 |||||||
 Db 2449 ACGGTCAACGCTTGTCTGTAACGCGATGCCGGGACAGACAAGCCCGTCAAGGCGCGCTCA 2508
 |||||||

QY 961 gcggtgttgcggggtgtcggggtcgtgcttaactatgagc 1001
 |||||||

Db 2509 GCGGCTGTTGGCGGCTGTGGGCTTGAATATGCGGC 2549

RESULT 2
 US-08-371-320-1
 ; Sequence 1, Application US/08371320
 ; Patent No. 5702888
 ; GENERAL INFORMATION:
 ; APPLICANT: Holte, Hans Joachim
 ; APPLICANT: Seidl, Rudolf
 ; APPLICANT: Schmitz, Gudrun
 ; APPLICANT: Scholer, Hans
 ; APPLICANT: Kessler, Christoph
 ; APPLICANT: Maties, Ralf
 ; TITLE OF INVENTION: Process for the Detection of Nucleic Acids
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Felte & Lynch
 ; STREET: 805 Third Avenue
 ; CITY: New York City
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10022
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: PC/DOS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/371,320
 ; FILING DATE: 11-JAN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/415,307
 ; FILING DATE: 09-Jan-1990
 ; APPLICATION NUMBER: PCT/EP89/0026
 ; FILING DATE: 12-Jan-1989
 ; APPLICATION NUMBER: DE 38 00 642.1
 ; FILING DATE: 12-Jan-1988
 ; APPLICATION NUMBER: DE 38 13 278.8
 ; FILING DATE: 20-Apr-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hanson, No. 5702888man D.
 ; REGISTRATION NUMBER: 30,946
 ; REFERENCE/DOCKET NUMBER: BOER 798
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-688-9200
 ; TELEFAX: 212-838-3884
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3104 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-371-320-1

Query Match 100.0%; Score 1001; DB 1; Length 3104;
 Best Local Similarity 100.0%; Pred. No. 2.1e-301;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttlatccgcctccatccagctcattatgtgtccgggaagctagaagtaagt 60
 |||||||
 Db 1549 CCTGCACACTTATCCGCGCTCATCCAGTATTAATTGTTCCGGGAGAGCTACAGTAAGT 1608
 |||||||

QY 61 agtgcgcagtaataagttgacgaacgttgttgcattgtctacaggaatcgttgtca 120
 |||||||
 Db 1609 AGTTCGCCAGTTAATAGTTGGCCCAACGTTGTGCAATTCCTACAGGCACTCGGTGCTCA 1668
 |||||||

QY 121 cgcctcgcgtttgttattggttattcagctcaggttcccaagatcaaggcgagttaca 180
 |||||||
 Db 1669 CGCTCGTCGTTGTGTAAGGCTTATTCAGCTCCGGTTCCTCCAGATCAAGGCGAGTTACA 1728
 |||||||

QY 181 tgatcccccatgtgtgcaaaaaaagcgtttagctcctcgtctccgactcgttgtcaga 240
 |||||||

Db	1729	TGATCCCATGTTGTGCAAAAAGGGTTACTCTCGGCTCCGATCGTTGTGAGA	1788
Oy	241	atgaagtgtgcgcagtgctatcaactatggtatgggaagcagtcataatcttcttact	300
Db	1789	AGTAAGTGGCCCACTGTTATCACTCAAGGTTATGGGAGACATGCATAATTCCTTACT	1848
Oy	301	gtatccacccgttaagatgcttctctgacttgg tgaactcaaaccaatcatctga	360
Db	1849	GTCATCGCATCCCTAAGATGCTTTTCTGTGACTGGTGTGATGCTCAACCAATCATCTGTA	1908
Oy	361	gaatagtgatagcgcgacagagttgtctctgtccgcgtcaatacggataataccgcy	420
Db	1909	GAATAGTGTGTGGCGACCGAGTGTCTCTTCCCGGGGTCAATACGGGATTAATACCGCG	1968
Oy	421	ccacatagcggaactttaaagtgctcatcatctggaacagctctctggcggaacac	480
Db	1969	CCAAATGCGAAGAACTTTAAAGTGCCTCATTGTGGAAAAAGCTTCTCGGGCGAAAACTC	2028
Oy	481	tcaagatcttaacgcgtgttgatgaltcagltgaltgaaccacactcgtgcacccaatga	540
Db	2029	TCAAGATCTTACCGCTGTGTGAATCATCACTGTGAATGAACCCACATCGTGCAACCAATGA	2088
Oy	541	tcttcagcatcttctactcttccacgcgcttctgggtgtagaanaaacggagaagcaaat	600
Db	2089	TCTTTCACCATCTTTTACTTTCACACACCGTTTCTGGGTGAGCAAAAACGAGAAAGCAAAAT	2148
Oy	601	gcgcgaataaaggaataaagcgacacggaatgttgatactatactcttcccttlt	660
Db	2149	GCGCCAAAAAAGGAAATMAAGCGCGACACGGAATGTGAATATCTATCTCTCTTTT	2208
Oy	661	caatatattgaagcatttaccaggttatgtgtcctcagtcagtcgagtcacatatattgaatgt	720
Db	2209	CAATATTATTGAAGCATTTATTCAGGGCTTATGTCTCATAGCGGATACATAATTGAAGT	2268
Oy	721	atgaagaataaatacaaaataggggtgtccgcgacattcccggaaggtgcacacttac	780
Db	2269	ATTTAGAAAAATTAACAATAAGGGGTTCCGCGCACATTTCCCGAAAAAGTGGCACTGTAC	2328
Oy	781	gtctaaagaacattatcatcatgacattaacctataaataagcgtatcaagaagccc	840
Db	2329	GTCTAAAGAAACATTATTATCATGACATTAACCTATTAATAAATAGCGATACGACGAGGCC	2388
Oy	841	tttcgtctcgcggttccgtgtgtatgcagtgtaaacctctgacaaatgcaatgagctccgag	900
Db	2389	TTTTGTCTCGGCGCTTTCGGGTGATGACGGGTAACCTTGACACATGCAAGCTCCCGAAG	2448
Oy	901	acggtaacagcttctgttaagcggatgcgcggagacagaacagccgctcaagcgcgctga	960
Db	2449	ACGGTCACAGCTTCTCTGTAAAGCGGAATCCCGGAGACAGAACGCCGTGAGGCGCGTCA	2508
Oy	961	gcggggtgttgcggggtgcggggctgtgcttaactatgcggc 1001	
Db	2509	GCGGGTGTGGCGGGTGTCCGGGCTGTGACTTAACATGATGCGGC 2549	
RESULT 3			
5453363-2/c			
Patent No. 5453363			
APPLICANT: RUDOLPH, RAINER; FISCHER, STEPHAN; MATTES, RALF			
TITLE OF INVENTION: PROCESS FOR THE ACTIVATION OF T-PA OR			
ING AFTER GENETIC EXPRESSION IN PROKARYOTES			
NUMBER OF SEQUENCES: 4			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/206,044			
FILING DATE: 02-MAR-1994			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 942,370			
FILING DATE: 09-SEP-1992			
APPLICATION NUMBER: 498,500			
FILING DATE: 23-MAR-1990			
APPLICATION NUMBER: 76,207			
FILING DATE: 23-OCT-1986			

SEQ ID NO:2:	LENGTH: 3343	5453363-2
Query Match	100.0%;	Score 1001; DB 6; Length 3343;
Best Local Similarity	100.0%;	Pred. No. 2.2e-301;
Matches 1001;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY 1	ccctgaactttatccgcctccatccagcctataattgttgcgggaagcctagatgaat	60
DB 2188	ccctgaactttatccgcctccatccagcctataattgttgcgggaagcctagatgaat	2129
QY 61	agttgcgaattaaatagtttgcgaacgttgttgcattgtgcacaggaatcgttggttca	120
DB 2128	agttgcgaattaaatagtttgcgaacgttgttgcattgtgcacaggaatcgttggttca	2069
QY 121	cgcctgcgttgtgtatggtcctcaatcagctccggttcccaacgatcaaggcagtttaca	180
DB 2068	cgcctgcgttgtgtatggtcctcaatcagctccggttcccaacgatcaaggcagtttaca	2009
QY 181	tgatcccccattgttgcgaaaaaacggttagctctcttgcgtccctccgattgttcaga	240
DB 2008	tgatcccccattgttgcgaaaaaacggttagctctcttgcgtccctccgattgttcaga	1949
QY 241	agtaagttgcgcagtgatcatcaatcattgttatgacagcactgtcataattcttact	300
DB 1948	agtaagttgcgcagtgatcatcaatcattgttatgacagcactgtcataattcttact	1889
QY 301	gtcatgtccatccgtaagaatgcttctgtgactgtgtgagtaactcaaccaatgtcttca	360
DB 1888	gtcatgtccatccgtaagaatgcttctgtgactgtgtgagtaactcaaccaatgtcttca	1829
QY 361	gaatggtgatgcgggacccaggttgtctcttgcggcgctcaatacgggataataccgcg	420
DB 1828	gaatggtgatgcgggacccaggttgtctcttgcggcgctcaatacgggataataccgcg	1769
QY 421	ccacatagacagaacttlaaagtgctcatcatctggaanaacgttcttcggggcganaaact	480
DB 1768	ccacatagacagaacttlaaagtgctcatcatctggaanaacgttcttcggggcganaaact	1709
QY 481	tcaagatcttaaccgctgttgtagatccagttcgatgtaaccacactcgttgacccaactga	540
DB 1708	tcaagatcttaaccgctgttgtagatccagttcgatgtaaccacactcgttgacccaactga	1649
QY 541	tcttgcgaactcttttactcttcaaccacggttttcgggttgagcaaaaaacggaaggaacaaat	600
DB 1648	tcttgcgaactcttttactcttcaaccacggttttcgggttgagcaaaaaacggaaggaacaaat	1589
QY 601	gccgcgaanaaagggaataagggcgacagcganaatgtgtlgaatactatactcttcctttt	660
DB 1588	gccgcgaanaaagggaataagggcgacagcganaatgtgtlgaatactatactcttcctttt	1529
QY 661	caatatattgaagcatltaacaggttattgtctcaatgacggagatacatatttgaatgt	720
DB 1528	caatatattgaagcatltaacaggttattgtctcaatgacggagatacatatttgaatgt	1469
QY 721	atttaaaaaataaaaaaataaggtttccgcgaacgttcccgaaaaagcgcaaccgtac	780
DB 1468	atttaaaaaataaaaaaataaggtttccgcgaacgttcccgaaaaagcgcaaccgtac	1409
QY 781	gtctaaagaaacattatcatcagcaatlaacattataaataagcgtlcatcagagacc	840
DB 1408	gtctaaagaaacattatcatcagcaatlaacattataaataagcgtlcatcagagacc	1349
QY 841	tttcgctcgcgcgtttcgggtgatgacggtgaaaacctcgcgaacatgacatgccgtccggag	900
DB 1348	tttcgctcgcgcgtttcgggtgatgacggtgaaaacctcgcgaacatgacatgccgtccggag	1289
QY 901	acggtaacacgttctcgtlgaacgagatgcgcggagacagaaacccgtcaaggcgctca	960
DB 1288	acggtaacacgttctcgtlgaacgagatgcgcggagacagaaacccgtcaaggcgctca	1229

QY 961 gccgagcttggcggggtcgcgggcttaactatgcggc 1001
|||||
Db 1228 gccggcttggcggggtcgcgggcttaactatgcggc 1188

RESULT 4

US-09-042-353-393
Sequence 393, Application US/09042353

Patent No. 6255458

GENERAL INFORMATION:

APPLICANT: Lomborg, Nils

APPLICANT: Kay, Robert M.

TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for

TITLE OF INVENTION: Producing Heterologous Antibodies

NUMBER OF SEQUENCES: 421

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIA TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/042,353

FILING DATE: 13-MAR-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/810,279

FILING DATE: 17-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/853,408

FILING DATE: 18-MAR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/904,068

FILING DATE: 23-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/990,860

FILING DATE: 16-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/053,131

FILING DATE: 26-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/096,762

FILING DATE: 22-JUL-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/155,301

FILING DATE: 18-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/161,739

FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/165,699

FILING DATE: 10-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/209,741

FILING DATE: 09-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/352,322

FILING DATE: 07-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/544,404

FILING DATE: 10-OCT-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/728,463

FILING DATE: 10-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US96/16433

FILING DATE: 10-OCT-1996

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/758,417

FILING DATE: 02-DEC-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/21803

FILING DATE: 01-DEC-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 014643-009040US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 393:

SEQUENCE CHARACTERISTICS:

LENGTH: 3819 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-09-042-353-393

Query Match 100.0%; Score 1001; DB 4; Length 3819;

Best Local Similarity 100.0%; Pred. No. 2,3e-301;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccgcgaacttattccgcctccatccatctataattgttcgcygaagcagaagtaagt 60
Db 1156 cctgcaactttatccgctccatccatccatccatccatccatccatccatccatccat 1215
QY 61 agtcgcgaattatagttgcgaagcgtgtgtccatctacgaagcagcgtgtgtga 120
Db 1216 agtcgcgaattatagttgcgaagcgtgtgtccatctacgaagcagcgtgtgtga 1275
QY 121 cgcctgcgttctgtatgctcattcaagtcgcgttcccaagcagcagcagcagcagc 180
Db 1276 cgcctgcgttctgtatgctcattcaagtcgcgttcccaagcagcagcagcagcagc 1335
QY 181 tgcaccccaatgttgcgaagcagcagcagcagcagcagcagcagcagcagcagc 240
Db 1336 tgcaccccaatgttgcgaagcagcagcagcagcagcagcagcagcagcagcagc 1395
QY 241 agtaagttgcgcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
Db 1396 agtaagttgcgcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1455
QY 301 gtcacgcatccgtaagatgcttctgtgacgtgtgacgtgtgacgtgtgacgtgtgac 360
Db 1456 gtcacgcatccgtaagatgcttctgtgacgtgtgacgtgtgacgtgtgacgtgtgac 1515
QY 361 gaataggtatgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420
Db 1516 gaataggtatgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1575
QY 421 ccaatagcagaacttaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
Db 1576 ccaatagcagaacttaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1635
QY 481 tcaagatcttaccgctgttgcagcagcagcagcagcagcagcagcagcagcagcagcagc 540
Db 1636 tcaagatcttaccgctgttgcagcagcagcagcagcagcagcagcagcagcagcagcagc 1695
QY 541 tcttgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 600
Db 1696 tcttgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1755
QY 601 gccgcaaaaaaggaataagggcagcagcagcagcagcagcagcagcagcagcagcagcagc 660
Db 1756 gccgcaaaaaaggaataagggcagcagcagcagcagcagcagcagcagcagcagcagcagc 1815
QY 661 caatatattgaagcattatcagcagcagcagcagcagcagcagcagcagcagcagcagc 720
|||||

Db 1936 GCTAAGAAACATATTATATCATGACATTAACCTATAAATAAGCGCTATACAGAGGCC 1995
Qy 841 ttctgtctgcgagcttccgttgatagcagtgtaaaacctcgtacacatgcagctccggag 900
Db 1996 TTTGCTCTCGCGCTTTCGGTGTGATGACGGTGAACCTCTGACACATGACGCTCCGGAG 2055
Qy 901 accgtcacagcttctgttaagcagatgcgagagcagacaagcccgctcagggcgctca 960
Db 2056 ACGGTACAGAGCTTGTCTGTAAGCGGATCGCGGAGCAGACAGAACCCGTCAGGGCGGTCA 2115
Qy 961 gcgggtgttgccgggtgtcggggctgcttaactatagggc 1001
Db 2116 GCGGGTGTGGCGGCTGCGGGCTGAGCTTAACATGCGGC 2156

RESULT 6
US-09-982A-21/c
; Sequence 21, Application US/0903982A
; Patent No. 625042
; GENERAL INFORMATION:
; APPLICANT: Cal, Zeling
; APPLICANT: Sprent, Jonathan
; APPLICANT: Brunmark, Anders
; APPLICANT: Jackson, Michael
; APPLICANT: Peterson, Per A
; TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR ACTIVATION OF T-CELL
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Olson & Hiern, Ltd.
; STREET: 20 No. 6225042th Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,982A
; FILING DATE: 16-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Olson, Arne M.
; REGISTRATION NUMBER: 30,203
; REFERENCE/DOCKET NUMBER: T5R14710
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 580-1180
; TELEFAX: (312) 580-1189
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3875 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-09-039-982A-21

Query Match 100.0%; Score 1001; DB 4; Length 3875;
Best Local Similarity 100.0%; Pred. No. 2.4e-301;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cctgaacttatacgcctccatccagctatataatgttgcgggaagctagagtaagt 60
Db 2491 CCTGCAACTTATCCGCTCCATCCAGTCATTAATGTTGCCGGGAAGCTAGAGTAGT 2432
Qy 61 agttgcagtaataagttgtgcgaagcttctgcacatgctacagagcatcgtgtgtca 120
Db 2431 AGTTGCGCAGTTAATGTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGTGTCA 2372

Qy 121 cgtcgtgtgttgatagcttcacgtccgttcccaacgaatcaagcgagttaca 180
Db 2371 CCGTGTGCTTTGTATGAGCTTCATTCAGCTCCGGTCCCAACGATCAAGCGAGTGTACA 2312
Qy 181 tgatcccccatgtcttgcaaaaaagcgttaagctccttcggtccctcagatcgtgtcaga 240
Db 2311 TGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCGGTCCCTCCATGCTGTCTCA 2252
Qy 241 agtaagttggccgagtgtaactcaatggttaagcagacatgacataatcttact 300
Db 2251 AGTAAGTTGGCCGAGTGTATACATCATGTTATGACAGACACTGATTAATTCCTTACT 2192
Qy 301 gtcattgcattcgttaagatgcttctctgtgactggtgagtaactcaacaaatcattcga 360
Db 2191 GTCATTGCCATCCGTAAAGATGCTTCTGTGACTGCTGAGTACTCAACCAATGCTATTCTGA 2132
Qy 361 gaatagttatgcgagcagagtgctcttgcgagcggtcacaatacaggaataacgcg 420
Db 2131 GAATAGTGTATGCGCGAGCCAGAGTGTGCTTCCCGGCGTCAATACGGGATTAATACCGC 2072
Qy 421 ccacatacagaactttaaagtgctcaatcattggaacagtlctcggggcgaaatc 480
Db 2071 CCACATACAGAACTTTAAAGTCTCATTTGGAACGTTCTTCCGGGCGGAAACTC 2012
Qy 481 tcaagatctaccgcttgatagatcagatcagatgtaacccactcgtgcacccactga 540
Db 2011 TCAAGATCTTACCGCTGTGAGATCCAGTTGATGATACCCACTCGTGACCCCACTGA 1952
Qy 541 tcttcagatcttacttaccagcgttctctgtgtgagcaaaaaagagagcaaat 600
Db 1951 TCTTCAGCATCTTTACTTTACACAGGCTTCTGAGTACGAAAAACAGAGCAAAAT 1892
Qy 601 gccgcaaaaaaggaataagggcgacacggaatggtgaatactcattcttctt 660
Db 1891 GCCGCAAAAAAGGGAATTAAGGCGCACAGAAATGTCATATCACTCTTCTTTT 1832
Qy 661 caatattatgaagattatcagagttatgtctcagagcagatataattgaaatgt 720
Db 1831 CAATATTATTGACATTTATCAAGGTTATTTCTCATGACCGGATCATATTTTGAATGT 1772
Qy 721 attagaataatacaaalaggggttcgcacatcttcccgcaaaagtgccacgtac 780
Db 1771 ATTAGAATAATTAACAATATGAGGGTTCGCGCACATTTCCCGCAAAAGTGCCACGTGAC 1712
Qy 781 gctaaagaaccattatatacagacattacctaataaataaggtatcagagggcc 840
Db 1711 GCTAAGAAACCATTTATATCATGACATTTAACTTAATAATAGCGGTATCAGAGGCC 1652
Qy 841 ttctgtcgcgagcttcgtgagcagtgtaaaacctcgcacatcagcaggtcccgag 900
Db 1651 TTTGCTCTCGCGCTTTCGGTGTGATGACGGTGAACCTCTGACACATGACGCTCCGGAG 1592
Qy 901 accgtcacagcttctgttaagcagatgcgagagcagacaagcccgctcagggcgctca 960
Db 1591 ACGGTACAGAGCTTGTCTGTAAGCGGATCGCGGAGCAGACAGACCCGTCAGGGCGGTCA 1532
Qy 961 gcgggtgttgccgggtgtcggggctgcttaactatagggc 1001
Db 1531 GCGGGTGTGGCGGCTGCGGGCTGAGCTTAACATGCGGC 1491

RESULT 7
US-09-039-641-21/c
; Sequence 21, Application US/09039641
; Patent No. 6251627
; GENERAL INFORMATION:
; APPLICANT: Cal, Zeling
; APPLICANT: Sprent, Jonathan
; APPLICANT: Brunmark, Anders
; APPLICANT: Jackson, Michael
; APPLICANT: Peterson, Per A
; TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR

TITLE OF INVENTION: ACTIVATION OF T-CELLS
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Olsson & Hierl, Ltd.
 STREET: 20 No. 6251627th Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/039,641
 FILING DATE: 8-MAR-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Olsson, Arne M.
 REGISTRATION NUMBER: 30, 203
 REFERENCE/DOCKET NUMBER: TSI14710
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 580-1180
 TELEFAX: (312) 580-1189
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3875 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-09-039-641-21

Query Match	100.0%	Score 1001	DB 4	Length 3875
Best Local Similarity	100.0%	Pred. No. 2.4e-301		
Matches 1001	Conservative	0	Mismatches 0	Indels 0
Oy	1	ccctgaacttaccgcgcctccatccagctcatatattgtgttcgsgaagctagaagta	60	
Db	2491	CTGTGCAACTTTATCCGCGCTCCATCCAGTCTATTAATTGTTCCGGGAAGCTAGAGTA	2432	
Oy	61	agttcgccagttlaataagtttgcgcgaagcttltgtgccattgtctaaaggaatcgtgtgtca	120	
Db	2431	AGTTGCGCAGTTAATAGTTTGGCGAACGTTGTGGCATTTGCTACAGGACATCGTGTCTCA	2372	
Oy	121	cgcctcgtcgttgtgtatggtcttcattcagctccggttccccaagataagcgagttaca	180	
Db	2371	CGCTCGTCGTTTGTTGGTTAGCTTCATTCAGCTCCGGTTCACCAAGCATCAAGCGCAGTTACA	2312	
Oy	181	tgtatcccccaatgttgcgaaaaaagcggttgaactcttcgcgttcctccgactcgtgtgtaca	240	
Db	2311	TGATGCCCCCATGTTGTGCAAAAAAGGGTTAGCTCTCTTCGGGTCCTCCGATGTTGTGAGA	2252	
Oy	241	agtaagttgcgcgcagttgttatacactcatgtaatgcgaacatgcataatcttcttact	300	
Db	2251	AGTAAGTTGGCGCGCATAGTTATCAGTCATCAGTGTATGSGAGACATGCAATATCTCTTACT	2192	
Oy	301	gtcatgccatcgtlaagatgtcttctctgtactgtgtgaatcaaccaagtaattctga	360	
Db	2191	GTCATGCCATCCCTAAGATAGCTTTTCTGTGACTGTGTGACTCTCAACCAAGTCACTTCTGA	2132	
Oy	361	gaataagtgatgttgcggcgagacgaattgtctcttgcgcggcggttaataagggataataccgg	420	
Db	2131	GAATAGTGTATGCGCGACCGGAGATGTCTCTTCCCGCGGCTCAATACGGGATAATACCGCG	2072	
Oy	421	cccatatagagaactttaaaagtgtccatcatcttggaaaaagcttctcggggcgaaatc	480	
Db	2071	CCCATATGCGAAGACTTTAAATAGTCTCATTCATTCGAAAAACGTTCTTCGGGCGAAAAATC	2012	

QY	481	tcgaagctcttacgcgctgtttgagatccaaagtcatgcatgtaaaccaactcgttcaccaacatga	540
Db	2011	TCGAAGACTCTTACCGCGTGTGAGATCCAGTTCGATTAACCCACTCGTGCACCCAACTGA	1955
QY	541	tcttcagcatccttctactcttcaccagcgtttctcgtgtgagcaaaaacaggaagcaaat	600
Db	1951	TCTTTCACGACTCTTCTTACTTTCACACAGGCTTTCGGGTGAGCAAAAACAGGAAGCAAAAT	1892
QY	601	gcgcgcaaaaaagggaaataaaggcgcaacgaaatgttgtatactactactcttcctttt	660
Db	1891	GCGCCAAAAAAGGGAATTAAGGGGACACGGAATGTGTGAATACTGTACTCTTCCTTTT	1833
QY	661	caaatattatgaagcatttcaaggyttattgtctcatgagcggataacatattgaaatgt	720
Db	1831	CAATATTATTGAAACATTTTATCAGAGGTTATTGTCATGAGCGGATACATATTTTGAATGT	1772
QY	721	attagaanaataaacaaatagaaggtttcgcgcgaactttcccgaaanaatgtccactgac	780
Db	1771	ATTATGAAAAAATTAACAAATAAGGGGTTCGCGGCACATTTTCCCGAAAAAGTCCACCTGAC	1711
QY	781	gtctaaagaacattatctatcatgacatllaacctataaanaatagaagcgtlcaacagagccc	840
Db	1711	GTCTAAGAAACATTATTATCATGACATTAACTATAAAAATAGGCGTATCAGAGGCC	1652
QY	841	tttcgctcgcgcgcttcgcgtgtatgacgftgaanaacctcgaacaatgtagctcccgagag	900
Db	1651	TTTGTGCTCGCGCGCTTTCGGGTGATGAGGGGTGAANAACCTCTGACACATGCAAGCTCCGGAG	1592
QY	901	acagtcacagacttctctgtaagcggatcgcgcggagagagaanaagcccgtcaaggcgcgtlca	960
Db	1591	ACGGTCACAGCTTGTCTGTAAAGCGAGATGCGGAGAGACAAACCCGTCAGGGCGGCTCA	1532
QY	961	gcggagtgtagcgcggtgtgcgggacttgacttaacatagcgc	1001
Db	1531	GCGGAGTGTGCGGGGTGTGCGGCTTGACTTAACATATCGCGC	1491

RESULT 8
US-09-039-762A-21/c
Sequence 21. Application US/09039762A
Patent No. 6255073
GENERAL INFORMATION:
APPLICANT: Cai, Zeling
APPLICANT: Sprent, Jonathan
APPLICANT: Brunmark, Anders
APPLICANT: Jackson, Michael
APPLICANT: Peterson, Pet A.
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS
TITLE OF INVENTION: FOR ACTIVATION OF T-CELLS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Olsson & Hierl, Ltd.
STREET: 20 No. 6255073th Wacker Drive, 36th Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,762A
FILING DATE: 16-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OLSON, Arne M.
REGISTRATION NUMBER: 30,203
REFERENCE/DOCKET NUMBER: TSRI 471.0 DIV.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 580-1180
TELEFAX: (312) 580-1189

; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3875 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; US-09-039-762A-21

Query Match 100.0%; Score 1001; DB 4; Length 3875;
 Best Local Similarity 100.0%; Pred. No. 2.4e-301;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaacttattccgctccatccagttcattatattgttccgggaagctagagtaagt 60
 DB 2491 CCTGCAACTTTATCCCGCCATCCAGTCTATTATTTGCGGGAAGCTAGAGTAAGT 2432
 QY 61 agtgcagtaataagtttgcaacggtttgtccattgtcagaggaagcgtgtgtca 120
 DB 2431 AGTTGCCAGTTAATAGTTGGCAAGGTTGTGCCATTGTCTACAGGCAATCGTGGTGCA 2372
 QY 121 cgtctcgtttgtatgttccatccatccagctccggttcccaagatcaagcgagttaca 180
 DB 2371 CGCTCGCTGTTTGATGTGGCTTCATCCAGTCCGGTCCCAACGATCAAGGAGTTACA 2312
 QY 181 tgatcccccagtgltgtcaaaaaaagcggttaagctccctcgtctccgagtcgttgcaga 240
 DB 2311 TGATCCCGCATGTTGTGCAAAAAAGGCGTTAGCTCTCCGTCCTCCGATCTGTGTCAGA 2252
 QY 241 agtaagtttgccgagtgatcatcactcagttatgagcagacacatcaatcttact 300
 DB 2251 AGTAAGTTGGCGCATGTTATCTACTCATGTATGTGGCAGCACTCATATATCTTACT 2192
 QY 301 gtcataccatccgtaagatgcttctgtgactgtgtgaactcaaccaagctattcga 360
 DB 2191 GTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGTAATCAACCAAGCATTTCTGA 2132
 QY 361 gaatagtgatgcgagcagaggtgtctctgtccggcgctcaatacggagataacgcg 420
 DB 2131 GAATAGTGTATCGCGGACGCGAGTGTCTCTTCCCGCGCTCAATCGGATATATCCCG 2072
 QY 421 ccaatagcagacactttaaagtgctcaatcattgaaacgttcttggggggaacatc 480
 DB 2071 CCACATAGCAGAACTTTAAAGTCTCATATTGAAAGCTTTCTTCCGGGGAACCTC 2012
 QY 481 tcaagatcttaccgctgtgagatccagttcagatgtaaccactcgtgcacccaactga 540
 DB 2011 TCAAGATCTTACCGCTGTGTGATCCAGTTGATGTAACCCACTCGTGACCCCAACGA 1952
 QY 541 tcttaagcacttcttaacttcaacagcgttctgtgtgagcaaaaaacggagacaaat 600
 DB 1951 TCTTACGATCTTTTACTTTCACACAGCGTTCTTGCGGTGAGCAAAAAACGAGGAAAT 1892
 QY 601 gccgcaaaaaaggaataaggcgagacagaaatgttgaatactaacttcttctt 660
 DB 1891 GCCGCAAAAAAGGATTAAGGCGACAGGAATTTGATTAATCTATCTTCTCTTTT 1832
 QY 661 caatatattgaagatattatcaggttattgtlcatcagagcgaataatattgaatg 720
 DB 1831 CAATATTATTGAAGCATTTATCAGGTTATTTCTCATGAGCGGATACATATTGTAAGT 1772
 QY 721 atttgaaaaaataaacaataaggggttccgcgacatttcccgaaaaagtccacgtac 780
 DB 1771 ATTGTAAAAAATAAACAAATAGGGGTTCCGCGCATTTTCCCGAAAAAGTCCACCTGAC 1712
 QY 781 gcttaagaacacattatcatcagcatctaataaataagacgcatcacagagccc 840
 DB 1711 GCTTAAGAAACCATTTATTTATCTGACATTTAACTTAAATAATAGCGATATACAGAGGCC 1652
 QY 841 ttctgctcgcgcttgcgtgatgacgltgaaaaactctgacacatgcagctcccgag 900

DB 1651 TTTCGTCGCGCGCTTTCGATGATGAGTGAAACCTTGACATGACAGCTCCCGAG 1592
 QY 901 acggtcaacgttctgtctgaagcgtatgcggagcagacaaagccgttcaggcgctca 960
 DB 1591 ACGGTCAACGCTTGTCTGTATGACGAGATGCCGAGACAGAACGCCCTGACGGCGCTCA 1532
 QY 961 gcgggtgttgcggggtgtcgcgggctggttaactatgcgc 1001
 DB 1531 GCGGGTGTGGCGGGGTGTGCGGGCTGCTTACTATATCGGC 1491

RESULT 9

US-09-039-982A-27/c
 Sequence 27, Application US/09039982A
 Patent No. 6225042

GENERAL INFORMATION:

APPLICANT: Cal, Zeling
 APPLICANT: Sprent, Jonathan
 APPLICANT: Brunmark, Anders
 APPLICANT: Jackson, Michael
 APPLICANT: Peterson, Per A

TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR ACTIVATION OF T-C
 NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Olson & Hierl, Ltd.
 STREET: 20 No. 6225042th Wacker Drive, Suite 3000
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/039,982A
 FILING DATE: 16-MAR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Olson, Arne M.
 REGISTRATION NUMBER: 30,203
 REFERENCE/DOCKET NUMBER: TSRI4710
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 580-1180
 TELEFAX: (312) 580-1189

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:
 LENGTH: 3878 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-09-039-982A-27

Query Match

Best Local Similarity 100.0%; Score 1001; DB 4; Length 3878;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaacttattccgctccatccagttcattatattgttccgggaagctagagtaagt 60
 DB 2494 CCTGCAACTTTATCCCGCTCCATCCAGTCTATTATTTGCGGGAAGCTAGAGTAAGT 2435
 QY 61 agtgcagtaataagtttgcaacggtttgtccattgtcagagcagcagcgtgtgtca 120
 DB 2434 AGTTGCCAGTTAATAGTTGGCAAGGTTGTGCCATTGTCTACAGGCAATCGTGGTGCA 2375
 QY 121 cgtctcgtttgtatgttccatccatccagctccggttcccaagatcaagcgagttaca 180

Db 2374 CGCTGCTGTTGGTATGGCTTCATTCAGCTCCGGTTCCACAGATCAAGCGAGTTACA 2315
Qy 181 tgatccccatgttgcacaaaaaagggtagctctcgtcgtccgcacgtgtgcaga 240
Db 2314 TGATCCCCCATGTTGTCAAAAAAGCGGTTAGCTCCTCGGTCCTCGATCGTTGTGAGA 2255
Qy 241 agtaagtgcgcgaggtgtatcaatagttagtgcaacgactgacatcttctact 300
Db 2254 AGTAAGTTGGCGCGAGTTATCATCTATGTTATGGACACACTGCAATAATCTCTTACT 2195
Qy 301 gtcatgcacatccgtaagatccttctcgtacgtgtgtactcaacaaagtcattctga 360
Db 2194 GTCATGCCATCCGTAACATGTTTCTGTGACGTGTGATGACACCAAGTCACTTCTGA 2135
Qy 361 gaatagtgtatgcgcgacgcagagttgctcttgcgcgcgtcaatacgggataaccg 420
Db 2134 GAATAGTGTAGCGCGACGACGAGTTGCTCTGCGCGGCTCAATACGGGATATACCGCG 2075
Qy 421 ccacatagcgaactttaaagtgctatcatctgtgaaacgcttctgcggcgcaaaactc 480
Db 2074 CCACATAGCAACATTTAAAGTCTCATCTATGTAAGAAAGCTTCTTGGGCGCAAACTC 2015
Qy 481 tcaagatcttaccgcgtgtgtatgcacgtgtatgcacacacacacacacacacacac 540
Db 2014 TCAAGGATCTTACCGCTGTGTGAGATCCAGTTCCATGTAACCCACTCGTGACCCAACTGA 1955
Qy 541 tcttcagatcttcttacttctcaccagcgttctgtgtgtgacaaaaaaggaagcaaat 600
Db 1954 TCTTCAGATCTTCTTACTTTCACACGAGGTTCTGTGTGAGCAAAAGAGCAAAAT 1895
Qy 601 gccgcaaaaaaggaataaagggcgacacgcagaaatggtgaactacacacacacacacac 660
Db 1894 GCCGCAAAAAAGGAATRAAGGCGCACACGGAATGTTGAATACATACATCTCTCTTTT 1835
Qy 661 caataattatgaacattatcaaggttatgtctcatgtagcgagatacataattgaatgt 720
Db 1834 CAATATTTATGACATTTATTCAGGTTATGTCTCATGACGAGTACATATTTGAATGT 1775
Qy 721 attagaataaataaataataggggttcgcgcacacacacacacacacacacacacac 780
Db 1774 ATTTAGAAAAATAAACAATAGGGGTTCCGCGCACATTTCCCAAAAGTCCACCTGAC 1715
Qy 781 gtctaaagaacattatcatatcatatcaatcaataaataaataaataaataaataaataa 840
Db 1714 GTCTAAGAAACATTTATTTATCATGACATTTAACTATTAATAATAGCGTATCAGAGGCC 1655
Qy 841 ttctgcctgcgcgttctgcgtatgcaggtgtgaaacacacacacacacacacacacac 900
Db 1654 TTTGCTGTCGGCGGCTTTGCGATGAGCGGTGAATAACCTGTGACACATGACACTCCGAG 1595
Qy 901 acggtcacagcttctgtatgaagcgatgcgcgagcgagacgaacacacacacacacac 960
Db 1594 ACGGTCAACAGCTTGTGTGTAAGCGGATGCCGCGAGACAGAACCGCTCAAGGCGGCTCA 1535
Qy 961 gcgggtgtgcgcgggtgtgcggggtgcgttgaactgcgcg 1001
Db 1534 GCGGSGTGTGGCGGGTGTGCGGGCTGGCTTAATATGCGGC 1494

RESULT 10
US-09-039-641-27/c

Sequence 27, Application US/09039641

GENERAL INFORMATION:

APPLICANT: Cal, Zelling

APPLICANT: Sprent, Jonathan

APPLICANT: Brunmark, Anders

APPLICANT: Jackson, Michael

APPLICANT: Peterson, Per A

TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Olson & Hierl, Ltd.
STREET: 20 No. 6251627th Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,641
FILING DATE: 8-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Olson, Arne M.
REGISTRATION NUMBER: 30, 203
REFERENCE/DOCKET NUMBER: TSR14710
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 580-1180
TELEFAX: (312) 580-1189
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3878 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-039-641-27

Query Match 100.0%; Score 1001; DB 4; Length 3878;
Best Local Similarity 100.0%; Pred. No. 2.4e-301;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cctgcacattatccgcctccatccagctctatattgttgcgcggaagctaggaagt 60
Db 2494 CCGCAACTTATCCCGCTCCATCCAGTCTATTATTTGTCGGGAACTAGAGTAAT 2435
Qy 61 agtgcacagtaaatagttgcgcaagttgttgcacatgtctacaggaactgtgtgtca 120
Db 2434 AGTTCGCATTTATATGTTTGGCAAGTTGTCATTTGCTACAGACGACMGCGTGCTCA 2375
Qy 121 cgtcgtcgttgtgtatgtgttcattcaagctccggttcccaacgatacaaggagttaca 180
Db 2374 CGCTCGTGTGTTGTTAGGCTTCATTCAGCTCCGTTCCCAACGATCAAGCGCAGTTACA 2315
Qy 181 tgatccccatgttgcacaaaaaagggtagctctcgtcgtccgcacgtgtgcaga 240
Db 2314 TGATCCCCCATGTTGTCAAAAAAGCGGTTAGCTCCTCGGTCCTCGATCGTTGTGAGA 2255
Qy 241 agtaagtgcgcgaggtgtatcaatagttagtgcaacgactgacatcttctact 300
Db 2254 AGTAAGTTGGCGCGAGTTATCATCTATGTTATGGACACACTGCAATAATCTCTTACT 2195
Qy 301 gtcatgcacatccgtaagatccttctcgtacgtgtgtactcaacaaagtcattctga 360
Db 2194 GTCATGCCATCCGTAACATGTTTCTGTGACGTGTGATGACACCAAGTCACTTCTGA 2135
Qy 361 gaatagtgtatgcgcgacgcagagttgctcttgcgcgcgtcaatacgggataaccg 420
Db 2134 GAATAGTGTAGCGCGACGACGAGTTGCTCTGCGCGGCTCAATACGGGATATACCGCG 2075
Qy 421 ccacatagcgaactttaaagtgctatcatctgtgaaacgcttctgcggcgcaaaactc 480
Db 2074 CCACATAGCAACATTTAAAGTCTCATCTATGTAAGAAAGCTTCTTGGGCGCAAACTC 2015
Qy 481 tcaagatcttaccgcgtgtgtatgcacgtgtatgcacacacacacacacacacacac 540
Db 2014 TCAAGGATCTTACCGCTGTGTGAGATCCAGTTCCATGTAACCCACTCGTGACCCAACTGA 1955

```

QY 541 tcttcagcatcttcttacttaccagcgtttcgtggtgagcaaaacaggaagcaaat 600
   |||
Db 1954 TCTTCAGCATCTTTTACTTTCACACGCTTTCTGGGTGAGCAAAACAGAGGCAAAAT 1895
QY 601 gcccaaaaaaaggaaataaggcgacacggaatgttgaatactacttcttctt 660
   |||
Db 1894 GCGCAAAAAAGGAAATAGGGGCGACACGAAATGTTGAATACTCTCTCTTTT 1835
QY 661 caatatattgaagcatcttaccaggttatcttctcagagcgatataattgaagt 720
   |||
Db 1834 CAATATTATTGAAGCATTTATACGGGTATTGTCTCATGAGGGAATCATATTGAATAT 1775
QY 721 attagaataataaacaataagaggttcgcgcgacattcccgaaagtgcacatgac 780
   |||
Db 1774 ATTGAATAAATAACAAATAGGGGTTCCGGCACATTTCCCGAAAGTGGCACCTGAC 1715
QY 781 gtctaagaacattatcatgacataactataaataagcgatatacagaagccc 840
   |||
Db 1714 GTCTAAGAAACCATTTATATCATGACATTAACTATAAATAAGGGTATACAGAGGCC 1655
QY 841 ttctcgttcgcggtttcgtgatgacggttgaataacctctgacatgagtcgccgag 900
   |||
Db 1654 TTTCTGCTCGCGGTTTGGTGTGATGACGGGTGAACCTCTGACACATGACGCTCCGGAG 1595
QY 901 acggtacagctgtctgtgaagcgatgcccggagacagaagcccgtaagcgctga 960
   |||
Db 1594 ACGGTACAGCTTTGTCTGTAACGGGATGCCGGGAGACAGACCCCGTACAGGGCGCTCA 1535
QY 961 gcgggtgttgcggggtgtcggggtgtgcttaactatgctgcgc 1001
   |||
Db 1534 GCGGGTGTGGGCGGTGTGCGGGCTGTGCTTAACATGATGCGGC 1494

```

RESULT 11

US-09-039-762A-27/c
Sequence 27, Application US/09039762A

Patent No. 6255073

GENERAL INFORMATION:

APPLICANT: Cal, Zeling

APPLICANT: Sprent, Jonathan

APPLICANT: Brunmark, Anders

APPLICANT: Jackson, Michael

APPLICANT: Peterson, Per A.

TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: Olson & Hierl, Ltd.

STREET: 20 No. 6255073th Wacker Drive, 36th Floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/039,762A

FILING DATE: 16-MAR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OLSON, Arne M.

REGISTRATION NUMBER: 30,203

REFERENCE/DOCKET NUMBER: TSRI 471.0 DIV.2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 580-1180

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 3878 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
us-09-039-762A-27

```

Query Match 100.0%; Score 1001; DB 4; Length 3878;
Best Local Similarity 100.0%; Pred. No. 2,4e-301;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacattatccgcctccatccagctcattatattgttgcgggaagctagaagat 60
Db 2494 CCTGCACTTTATCCGCTCCATCATCTATTATTTGTTCCGGGAAGCTAGAGTAACT 2435
QY 61 agtcgcagttaaagtttgcgaacgtttgtccaatgtctacaagcgatcgttgtca 120
Db 2434 AGTTGCGCAGTTAATAGTTTGCGCAAGGTTTGCCATTGCTACAGGCATCGTGTGTA 2375
QY 121 cgtcgttcgtttgtatagcttcaatcagctcgggtcccaagatcaagcgagttaca 180
Db 2374 CGCTCGTCTGTTGATGCGCTTCATTCAGCTCCGCTTCCCAACATCAAGCGCAGTTACA 2315
QY 181 tgatcccccattgtgttgcataaaagcggttaagctccttcgttcctcgatcgttgtcaga 240
Db 2314 TGATCCCCCATGTTGTGCAAAAAAGCGGTTTACTCTTCGCTCTCCGATCGTTGTGACA 2255
QY 241 agtaagttgacgagttatcatcatatgttataagcaagctatcaatcttctact 300
Db 2254 AGTAAGTTGGCGGAGTGTATCATCTCATGATGTTATGGCAGACATGCATTAATTCCTTACT 2195
QY 301 gtcattgcacatccgtaagatcttctcgtgacgtgtgagtaacccaagaagtcattcga 360
Db 2194 GTCATGCCATCCGTAAGATGCTTTCTGTGACGTGTGATCAACCAAGTCAATTCCTGA 2135
QY 361 gaatagttatgctgcgcgacgagttgtctctgcgcgcgctcaatacgggataaccg 420
Db 2134 GAATAGTATAGCGCGGACCGACGAGTGTCTTCCCGGCGCTCATACGGGATTAATACCGG 2075
QY 421 ccacatagcagaacttaaaagtgtcatcatatggaanaacgtcttcggcgcaaaatc 480
Db 2074 CCACATAGCAAACTTTAAAGTCTCATCATGGAAGAGTTCTTCGGGCGCAAAATC 2015
QY 481 tcaaggtcttaccgctgttgaatcgaatcgaatgaaacccctgtgacccaactga 540
Db 2014 TCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTCGTGACCCCACTGA 1955
QY 541 tcttcagcatcttcttacttccacagcgtttctggtgagcaaaacaggaagcaaat 600
Db 1954 TCTTCAGCATCTTTTACTTTCACACGCTTTCTGGGTGAGCAAAACAGAGGCAAAAT 1895
QY 601 gcccaaaaaaaggaaataaggcgacacggaatgttgaatactacttcttctt 660
Db 1894 GCGCAAAAAAGGAAATAGGGGCGACACGAAATGTTGAATACTCTCTCTTTT 1835
QY 661 caatatattgaagcatcttaccaggttatcttctcagagcgatataattgaagt 720
Db 1834 CAATATTATTGAAGCATTTATACGGGTATTGTCTCATGAGGGAATCATATTGAATAT 1775
QY 721 attagaataataaacaataagaggttcgcgcgacattcccgaaagtgcacatgac 780
Db 1774 ATTGAATAAATAACAAATAGGGGTTCCGGCACATTTCCCGAAAGTGGCACCTGAC 1715
QY 781 gtctaagaacattatcatgacataactataaataagcgatatacagaagccc 840
Db 1714 GTCTAAGAAACCATTTATATCATGACATTAACTATAAATAAGGGTATACAGAGGCC 1655
QY 841 ttctcgttcgcggtttcgtgatgacggttgaataacctctgacatgagtcgccgag 900
Db 1654 TTTCTGCTCGCGGTTTGGTGTGATGACGGGTGAACCTCTGACACATGACGCTCCGGAG 1595

```

Oy 901 accgtacagctgtctgttaagcgatcgagcgagacagccctcagcgagctca 960
Db 1594 ACGGTCACAGCTTGTCTGTAAAGCGGATGCCGGAGCAGACAGCCCTCAGGGGGCTCA 1335
Oy 961 gccgggtgttgccgggtgttcgggggtctgagcttaactatcgagc 1001
Db 1534 GCGGGTGTGGGCGGTGTCCGGGCTGGCTTAACATATCGCGC 1494

RESULT 12

US-09-042-353-369
Sequence 369, Application US/09042353
Patent No. 6255458
GENERAL INFORMATION:
APPLICANT: Lomborg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 421
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,353
FILING DATE: 13-MAR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/810,279
FILING DATE: 17-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/853,408
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/904,068
FILING DATE: 23-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/096,762
FILING DATE: 22-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/155,301
FILING DATE: 18-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,739
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/165,699
FILING DATE: 10-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,741
FILING DATE: 09-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/728,463

FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO PCT/US96/16433
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,417
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO PCT/US97/21803
FILING DATE: 01-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 014643-009040US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 369:
SEQUENCE CHARACTERISTICS:
LENGTH: 3881 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-042-353-369

Query Match 100.0%; Score 1001; DB 4; Length 3881;
Best Local Similarity 100.0%; Pred. No. 2.4e-301;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cctgcacattatccgcctccatccagctcattatattgttcgaggaagctagtaagt 60
Db 1156 CCGCAACTTATTCGCCCTCCATCCAGCTATTAATTGTTCCGGGAGCTAGATAACT 1215
Oy 61 agttccagcttaataagatttgcgaacgttgcattgctacaggaatcgtgtgtca 120
Db 1216 AGTTCGCCAGTAAATAGTTGCGCAAGTTGTCATTGCTACAGGATCGGTGTCA 1275
Oy 121 cgcctcgttctgtatagcttattcagctccggttcccaagatcaaggagttaca 180
Db 1276 CGCTCGCTTGTGTATGCTTATTCAGCTCCGTTCCCAAGATCAAGCGAGTTACA 1335
Oy 181 tgaatcccatgttgcgaacaaagcggttagctccttcgctccgactcgtgtcaga 240
Db 1336 TGATCCCCCATGTGTGCAAAAAAGCGTTAGCTCCTCGCTCCTCGATCGTTGACAA 1395
Oy 241 agtaagttgcccagctgtatcaactcaactaagtgtatgagcagcactgataatcttact 300
Db 1396 AGTAAGTTGGCCGACAGTATCATCATCATGTTATGACAGCACTGCATTAATCTTACT 1455
Oy 301 gtaatgccatccgtaagatcttctgtgactggtgagttactcaaccgaattattcga 360
Db 1456 GTCAATGCCATCCGTAATATCTTTCTGTGACGTGTGAGTACTCAACCAAGTATTCTGA 1515
Oy 361 gaatagttatgcgagccaggttgccttcgcccgcgcgtcaatacaggaataacagcg 420
Db 1516 GAATAGTGTATGCGGCGACCGAGTGTCTTGTCCCGCGCTCATAGGATATATACGGG 1575
Oy 421 ccaatagcagaacattaaagtctcaatcaatgtgaacacgttcttcggggcgaaactc 480
Db 1576 CCACATAGCAGAACTTAAAGTGCATCATATGGAAGAGTTCTTCCGGGCGCAAAATC 1635
Oy 481 tcaagatcttaccgctggtgagatccagttgatgataaccacactgtgacccaactga 540
Db 1636 TCAAGGATCTTACCGCTGTGAGATCCAGTTGATGATCAACCACTGTGACCACTGA 1695
Oy 541 tcttcagcatttacttaccacagcgttctggtgtgagcaaaaaacaggaaggaanaat 600
Db 1696 TCTTCACATCTTTTACTTTCACCAAGGTTCTGGGTGAGCAAAAAACAGGAAGCAAAAT 1755
Oy 601 gccgcaaaaaaggaataaggcgcaacgggaatgtgaatactactcttcttctt 660


```

APPLICANT: Brunmark, Anders
APPLICANT: Jackson, Michael
APPLICANT: Peterson, Per A
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR
TITLE OF INVENTION: ACTIVATION OF T-CELLS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: Olson & Hierl, Ltd.
STREET: 20 No. 6251627th Wacker Drive, Suite 3000
City: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,641
FILING DATE: 8-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Olson, Arne M.
REGISTRATION NUMBER: 30,203
REFERENCE/DOCKET NUMBER: TSRI4710
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 580-1180
TELEFAX: (312) 580-1189
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 3883 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-039-641-30

```

Query Match 100.0%; Score 1001; DB 4; Length 3883;
 Best Local Similarity 100.0%; Pred. No. 2.4e-301;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgcacacttaccgcccacatccagctctattatgttgcgggaagctagagtaagt 60
DB 2499 CCTGCACACTTTATCCGCTCCATCCAGTCTATTATGTTGCCGGGAAGCTAGAGTAAGT 2440
QY 61 agtcgcagatlaaatagtttgcgaacgtttgtccatgtctacaggaatcgtgtgtca 120
DB 2439 AGTTCGCCAGTTAATAGTTTGCGCAACGTTGTGCCATTGCTACAGGCATCGTGGTCTCA 2380
QY 121 cgtctcgttcttgatagcttcaatcagctccggttcccaagaatcaagcgagttaca 180
DB 2379 CGCTCGTCTGTTGTATGCTTCATTCACCTCGGTTCCCAAGATCAAGCGAGTTACA 2320
QY 181 tcatcccatcttctgcaaaaagcggttaactcttcgtctcgtcgtatgttgcaga 240
DB 2319 TCAATCCCATGTTGTGCAAAAAGGGTTAGCTCTCTGCTCCGATCCGATCGTGTGAGA 2260
QY 241 agtaagtttgcgcagttgtatcaactatggtatgacagacactgataatcttctact 300
DB 2259 AGTAAGTTGGCGGAGTGTATACATCATGTTATGGCAGCAGCTGCATTAATCTTACT 2200
QY 301 gtcaatgcacatccgtaagaatgtcttctgtgactgtgtgagtaactcaacaagtaacttca 360
DB 2199 GTCAATGCCATCCGTAAAGATGCTTTCTGTGACTGGTGAATCAACCAAGTCAATCTGA 2140
QY 361 gaatagttatgcgcagacaggttgccttgcggcggtacataagggataatccggc 420
DB 2139 GAATAGTTATGCGGCGAGCGAGTTGCTCTTCCGCGGCTCAATAGCGATTAATACCGCG 2080

```

```

QY 421 ccacataagaacactttaaagttgcatcacttgaaacgcttctcgggcgaataatc 480
DB 2079 CCACATAGCAGACACTTTAAAGTGTCTATCATTTGGAATAAACTTCTCGGGCGAAAACTC 2020
QY 481 tcaagatcttaccgcttcttgatccagttcagatgtaaaccaactcgtgcaccaactga 540
DB 2019 TCAAGATCTTACCGCTGTGATCCAGTTGATGTAACTGACCTGTCACCAACTGA 1960
QY 541 tcttaagatcttacttaccagcggttctcgtgtgagtaaaaagaggaagcaaat 600
DB 1959 TCTTACGATCTTTTATCTTTCACACGCGTTCTGGGTGACAAAAACGAAAGCAAAAT 1900
QY 601 gccgcaaaaagaggaataagggcgacagaaatgttgaatactcatacttcttctt 660
DB 1899 GCCGCAAAAAGGGAATAGGGCGACAGCAAAATGTTGAATACATACATCTTCTTTT 1840
QY 661 caatattctgaagcattatcaaggttatttgcctcalgagcgagatacataattgaatgt 720
DB 1839 CAATATTATTGAAGCATTTATCAGGGTTATGCTCATGACGGGATCATATTGGAATGT 1780
QY 721 attagaataataacaataaggggttcgcgcacattcccgaaaagtgccactgac 780
DB 1779 ATTAGAATAAATAAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTCCACCTGAC 1720
QY 781 gtctaagaacattatatacatgacattaaactataaaataagcgctatcacagagccc 840
DB 1719 GTCTAAGAACAACATATTATATCATGACATTAACTATAAAATFAGCGGTATCAGAGGCC 1660
QY 841 ttctgctcgcggttctgtgtatgaaggttgaaaaactctgacacatgacgtcccgag 900
DB 1659 TTTCTCTTCGCGCGCTTTCGGTGTGATGACGGTGAACCTCTGCACATGACAGCTCCCGAG 1600
QY 901 acgttcacagcttctgttgaagcgatgacggagcgagacagacccgtcggggcggtca 960
DB 1599 ACGGTACACAGCTTGTCTGTAAAGCGGATGCCGGAGCAGACAAAGCCCTGACAGCGCGCTCA 1540
QY 961 gcgggtgttgccgggtgtgcggggtgacttactaactatgcgcgc 1001
DB 1539 GCGGCTGTGGCGGGGTGTGCGGGCTGGCTTAATATGCGGC 1499

```

Search completed: January 17, 2002, 16:26:57
 Job time: 12327 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 14:58:38 ; Search time 3755.78 Seconds
(without alignments)
2863.992 Million cell updates/sec

Title: US-09-740-211-13_COPY_6000_7000

Perfect score: 1001

Sequence: 1 cctgaacttaccgcgcctc.....ggctgcttaactatgcgcgc 1001

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

EST:*

- 1: em_estfun:*
- 2: em_esthm:*
- 3: em_estin:*
- 4: em_estom:*
- 5: em_estpl:*
- 6: em_estba:*
- 7: em_estro:*
- 8: em_estov:*
- 9: em_hlc:*
- 10: gb_estl:*
- 11: gb_est2:*
- 12: gb_hlc:*
- 13: gb_gss:*
- 14: em_gss_fun:*
- 15: em_gss_hum:*
- 16: em_gss_inv:*
- 17: em_gss_pln:*
- 18: em_gss_pro:*
- 19: em_gss_rnd:*
- 20: em_gss_vit:*
- 21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	766.8	76.6	935	11	BG838279 GC01_10e0
C 2	741	74.0	841	10	AL042026
C 3	673.2	67.3	754	11	BG920379 DKF2P34E
C 4	650.6	65.0	767	10	AL040542
C 5	638	63.7	693	10	AL598527 DKF2P3130
C 6	636.6	63.6	688	10	AL042640
C 7	607.2	60.7	914	10	BE569261
C 8	591.4	59.1	702	10	AL042540
C 9	588.2	58.8	695	13	AG002183 Homo sapi
C 10	576.8	57.6	658	13	AG014918 Homo sapi
C 11	559.2	55.9	793	13	AG000364 Homo sapi
C 12	554	55.3	696	13	AG009765 Homo sapi

13	548.4	54.8	696	10	BE573230	BE573230 601333125
14	546.4	54.6	801	10	BE569687	BE569687 601332450
C 15	545.6	54.5	554	10	AL041808	AL041808 DKF2P34B
C 16	541.6	54.1	720	13	AG000534	AG000534 Homo sapi
C 17	531.8	53.1	711	13	AG010947	AG010947 Homo sapi
C 18	531.4	53.1	890	10	AL046135	AL046135 DKF2P34P
C 19	517.6	51.7	619	10	AV611386	AV611386 AV611386
C 20	508.6	50.8	710	13	CMS075M1	AL430631 clone BA0
C 21	499.6	49.9	575	13	AG218759	AG218759 1006006H0
C 22	496	49.6	620	10	AL599532	AL599532 DKF2P313D
C 23	488.6	48.8	604	13	AG002706	AG002706 Homo sapi
C 24	488	48.8	499	10	AL041936	AL041936 AV612734
C 25	476.6	47.6	551	10	AL041936	AL041936 DKF2P34L
C 26	467	46.7	683	11	BG838185	BG838185 GC01_06e0
C 27	466.2	46.6	890	13	BO7159	BO7159 cSR172-cl-u
C 28	461	46.1	721	10	AL044583	AL044583 DKF2P34N
C 29	459.8	45.9	725	13	AG001126	AG001126 Homo sapi
C 30	457	45.7	647	13	AG008943	AG008943 Homo sapi
C 31	453	45.3	832	11	BG923768	BG923768 602825893
C 32	451.4	45.1	474	10	AV601733	AV601733 AV601733
C 33	451.2	45.1	684	13	AG014348	AG014348 Homo sapi
C 34	440	44.0	911	11	BG923556	BG923556 602823195
C 35	435	43.5	706	13	AG000533	AG000533 Homo sapi
C 36	433.2	43.3	686	13	AG016243	AG016243 Homo sapi
C 37	429	42.9	705	13	AG004170	AG004170 Homo sapi
C 38	429	42.9	705	13	AG010296	AG010296 Homo sapi
C 39	423	42.3	715	13	AG013813	AG013813 Homo sapi
C 40	418.8	41.8	751	13	AG007414	AG007414 Homo sapi
C 41	414	41.4	414	10	AV404393	AV404393 AV404393
C 42	413.6	41.3	501	10	AV596377	AV596377 AV596377
C 43	410	41.0	718	13	AG009622	AG009622 Homo sapi
C 44	409	40.9	640	13	B84895	B84895 RPL11-28L6
C 45	406	40.6	738	10	AL598454	AL598454 DKF2P313H

ALIGNMENTS

RESULT 1
BG838279/c 935 bp mRNA EST 25-MAY-2001
GC01_10e07.R GC01_AAFc_ECORC_cold_stressed_glycine_clandestina
Glycine clandestina cDNA clone GC01_10e07, mRNA sequence.

ACCESSION BG838279 GI:14204601
VERSION BG838279.1
KEYWORDS EST.
SOURCE Glycine clandestina.
ORGANISM Glycine clandestina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 935)
AUTHORS Singh,J.A., Farrah,S., Chapados,J., Courroux,P., De Moors,A., Harris
L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Spott,D. and Tinker
N.A.

Expressed Sequence Tags from Cold-Stressed Glycine clandestina
Seedlings
Unpublished (2001)
Contact: Singh,J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
Location/Qualifiers

FEATURES
source
1..935
/organism="Glycine clandestina"
/cultivar="1035"
/db_xref="taxon:45687"
/clone="GC01_10e07"
/clone_1fb="GC01_AAFc_ECORC_cold_stressed_glycine_clandest

ina"
/note="Vector: Bluescript SK+/XhoI-RCORI; Site_1: EORI;
Site_2: XhoI; Plants incubated at 2 degrees under 12 hours
of light/day. Harvested after only 2-3 days of cold
treatment. cDNA was prepared with the Uni-Zap cDNA kit
from Stratagene. Eco RI adapters were linked followed
by digest with Xho I/Eco RI and ligated to pBluescript."
BASE COUNT 243 a 213 c 227 g 238 t 14 others
ORIGIN

Query Match 76.6%; Score 766.8; DB 11; Length 935;
Best Local Similarity 98.5%; Pred. No. 1.9e-209;
Matches 776; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

```

QY 1 cctgcgaacttaccgctccaccagctcattatgttgcgggaagctagaagt 60
DB 813 CCTGCACCTTATCCGCCYCCATCCATCATTAATTGTCGGGAGCTAGAGTAAGT 754
QY 61 agtcgcgaactaataattgacgaacgttgtgacattgacagacatcgtgtgtca 120
DB 753 AGTTCGCCCACTTAATATGTTGGCCACAGTGTGTCATTCAGGCAATCGTGTGTCA 694
QY 121 cgtcgtcgttggatgcttcaatcagctccggtcccaacgacatcaagcgaattaca 180
DB 693 CGCTCGTCGTTGGTATGCTTCATTCAGCTCGGTCCCAACGATCARGCGAGTTTACA 634
QY 181 tgatcccccactgtgtgcaaaaagcggttagctccttgcgtccgactcgtgtgcaaa 240
DB 633 TGATCCCCCATGTTGTGCAAAAAGGGTATGAGTCTCGTCCGCTCCGATCGTGTGCA 574
QY 241 agtaagtt-ggccgcagtggtatcactcagtgatagcagacatcattctcttac 299
DB 573 AGTAAGTTGGCGCGCAGTGTATCACTCATGTTATGCGACACATCATTAATCTCTTAC 514
QY 300 tgcacgacatccgtaagatgcttctgtgactggtgactcaaccaagtcattctg 359
DB 513 TGTTCATGCCATCGTAAGATGCTTTCTGTGATGCTGATCACTCAACCAAGCATATTCTG 454
QY 360 agaatagtgatgcgacgacgacgacgtgctcctgcccgcggtcaatcaggaataccgc 419
DB 453 AGAATAGTGATGCGCGCGACCGAGTCTCTTCCCGCGCTCAATACGGGATTAATACCGC 394
QY 420 gccacatagcaaaccttcaaaagtctcactcacttggaaaaagcttctcgaggcgaaact 479
DB 393 GCCACATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTCGGGGCGAAACT 334
QY 480 ctcaagacatcagcgtgttgagacagcttcagatgataccacatcgtgcacccaactg 539
DB 333 CTCAGGAGCTTACCGCTGTGAGATCCAGTTCGATGTAMCCACATCGTGCAACCACTG 274
QY 540 atcttcagacatcttacttcaacagcgttctggtgtagcaaaaaacggaagcaaaa 599
DB 273 ATCTTCAGCATCTTTTACCTTACACGAGCTTTCTGGGTAGCAAAAACGGAAGCAAAA 214
QY 600 tgcgcgaaaaaaaggagataaggcgacacggaatgtgataactcattcttcccttt 659
DB 213 TGCGCGAAAAAAGGAAATAGGGCGCACGGAATGTTAAATACATCACTCTCTCTTTT 154
QY 660 tcaatttattgaagcattatcagagttatgttctcaagacggaatcattatggaatg 719
DB 153 TCAATATTTATTGAAGCATTTATCAGGGTATTGTTCTCATGAGCGGATACATATTTTGAATG 94
QY 720 tatttagaaaaaataaacaatagaggttccgcgacatttcccgcaaaagtgcacactga 779
DB 93 TATTTAGAAAAAATAAACAATAGGGGTTCGGCGACATTTCCCGAAAAAGTCCCAACTAA 34
QY 780 cgtctcaag 787
DB 33 ATTGTAAG 26

```

RESULT 2
AL042026/c
LOCUS
DEFINITION
AL042026
ACCESSION
AL042026
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL042026 841 bp mRNA EST 29-FEB-2000
DKFZp434E11.1-r1 434 (synonym: htes3) Homo sapiens cDNA clone
DKFZp434E11.1 5', mRNA sequence.
AL042026
AL042026.1 GI:5421372
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 841)
Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
Unpublished (1999)
Contact: Poustka A.J.
Department Leirach
Max-Planck-Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center.
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No SI sequence available.
This clone (DKFZp434E11.1) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

location/Qualifiers
1..841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434E11.1"
/clone_id="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"
BASE COUNT 226 a 192 c 202 g 221 t
ORIGIN

Query Match 74.0%; Score 741; DB 10; Length 841;
Best Local Similarity 100.0%; Pred. No. 4.8e-202;
Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgcgaacttaccgctccaccagctcattatgttgcgggaagctagaagt 60
DB 741 CCTGCACCTTATCCGCCCTCATTCATCTATTAATTGTTCCGGGAAAGCTTAGAGTAAGT 662
QY 61 agtcgcgaactaataattgacgaacgttgtgacattgacagacatcgtgtgtca 120
DB 681 AGTTCGCCCACTTAATATGTTGGCCACGTTGTCATTCAGGCAATCGTGTGTCA 622
QY 121 cgtcgtcgttggatgcttcaatcagctccggtcccaacgacatcaagcgaattaca 180
DB 621 CGCTCGTCGTTGGTATGCTTCATTCAGCTCGGTCCCAACGATCAAGCGAGTTTAC 562
QY 181 tgatcccccactgtgtgcaaaaagcggttagctccttgcgtccgacatcgtgtcaga 240
DB 561 TGATCCCCCATGTTGTGCAAAAAGCGGTATAGCTCTTCGCTCTCGAGCTGTGTGCA 502
QY 241 agtaagttggcgcagtgatcactcactcagtgatagcagcagcagataatctctact 300
DB 501 AGTAAGTTGGCGCGCAGTGTATCACTCATGTTATGCGACACATCATTAATCTCTTACT 442
QY 301 gtcatgcatccgtaagaatgcttctgtgactggtgagatcactcaaccaagtcattctga 360

```

Db	Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	441	62826032F1	NCI_CGAP_Mam6	Mus musculus	CDNA clone	IMAGE:4954975				
Qy	361	gaatactgatacgagcagccaggtgctcttgcgcccgcgclcaatacaggaataacccg	420							
Db	381	gaattactgatatggcgagccaggttgccttgcgcccgcgclcaatacaggaataacccg	322							
Qy	421	ccacataagcgaacttaaaagtgctcatatggaaaagctcttcggggcgaaaactc	480							
Db	321	ccacataagcgaacttaaaagtgctcatatggaaaagctcttcggggcgaaaactc	262							
Qy	481	tcaagaatcttaccggtgtgtgagatccaggttcgatgtaaccactcgttgaccacaatga	540							
Db	261	tcaagaatcttaccggtgtgtgagatccaggttcgatgtaaccactcgttgaccacaatga	202							
Qy	541	tcttcacatcttcttacttccaccgagcttcttggtgtgagcaaaaacgaaggaacaaat	600							
Db	201	tcttcacatcttcttacttccaccgagcttcttggtgtgagcaaaaacgaaggaacaaat	142							
Qy	601	gccgcacaaaaggaataaggcgagacacgaagaatgttgaatactcatcttctctttt	660							
Db	141	gccgcacaaaaggaataaggcgagacacgaagaatgttgaatactcatcttctctttt	82							
Qy	661	caatattatgaagcattatcaagggtatatttccatgaagcgaatcatattgaatgt	720							
Db	81	caatattatgaagcattatcaagggtatatttccatgaagcgaatcatattgaatgt	22							
Qy	721	attagaataataacaata 741								
Db	21	attagaataataacaata 1								
RESULT	3									
LOCUS	BG920379	754 bp	mRNA	EST	05-JUN-2001					
DEFINITION	602826032F1 NCI_CGAP_Mam6	Mus musculus	CDNA clone	IMAGE:4954975	5'					
ACCESSION	BG920379									
VERSION	BG920379.1	GI:14300855								
KEYWORDS	EST.									
SOURCE	house mouse.									
ORGANISM	Mus musculus									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
TITLE	1 (bases 1 to 754)									
JOURNAL	NIH-MGC http://mgi.nci.nih.gov/									
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)									
	Unpublished (1999)									
	Contact: Robert Strausberg, Ph.D.									
	Email: cgabbs-r@mail.nih.gov									
	Tissue Procurement: Jeffrey Green M.D.									
	CDNA Library Preparation: Life Technologies, Inc.									
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)									
	DNA Sequencing by: Incyte Genomics, Inc.									
	Clone distribution: MGC clone distribution information can be									
	found through the I.M.A.G.E. Consortium/LLNL at:									
	http://image.llnl.gov									
	Plate: ILAM10916 row: 0 column: 08									
	High quality sequence start: 15									
	High quality sequence stop: 717.									
FEATURES	Location/Qualifiers									
SOURCE	1..754									
	/organism="Mus musculus"									
	/strain="FVB/N"									
	/db_xref="taxon:10090"									
	/clone="IMAGE:4954975"									
	/clone_1lb="NCI_CGAP_Mam6"									
	/sex="Female; virgin"									

BASE COUNT	201 a	174 c	166 g	213 t	providing samples: Jeffrey Green, M.D., NIH*
Query Match					67.3%; Score 673.2; DB 11; Length 754;
Best Local Similarity					96.5%; Pred. No. 1,5e-182;
Matches 710; Conservative					0; Mismatches 23; Indels 3; Gaps
OY	77	gtttgcgaacgattgtgttcacattgtctacagacatcgttggtgtcagcctcgtgtgtgtga	136		
Db	11	gctttggcgaacgattgttgccattgctctacagacatcgttggtgtcagcctcgtgtgtga	70		
OY	137	tggtcttcacgttcgcgtgttcccaagatccaaggcgagttacatgatccccatgtt	196		
Db	71	tggcttcaattacgttcgcgtgttcccaagatccaaggcgagttacatgatccccatgtt	130		
OY	197	gcacaaagcggttagatgccttcgttcctccgatcgtgttcgaagtaagtgtgcgcag	256		
Db	131	gcacaaagcggttagatgccttcgttcctccgatcgtgttcgaagtaagtgtgcgcag	190		
OY	257	tgcttcaactatgttattggaacgacatgataatctcttactgtcatatgcctcgtaa	316		
Db	191	tgtttatcactatggtattgacacacacgcataatttcttactgtcatatgcctcgtaa	250		
OY	317	gatgcttcttgtagactggtgagttactccaacgaagtatctctgaagtaagtgtatgcgc	376		
Db	251	gatgcttcttgtagactggtgagttactccaacgaagtatctctgaagtaagtgtatgcgc	310		
OY	377	gacggaattgtcctctgtccgcgcgtcaataacggaataaacacggaattgtacgacat	436		
Db	311	gacggaattgtcctctgtccgcgcgtcaataacggaattgtacgacat	370		
OY	437	taaaagtgtcatcatatgtgnaaaacgtttcttcggggcgaaaacttcaaggatcttaccg	496		
Db	371	taaaagtgtcatcatatgtgnaaaacgtttcttcggggcgaaaacttcaaggatcttaccg	430		
OY	497	tgttgaagatcagttcgtatgttaacccactcgttgacccaactgtgttcaagatcttcta	556		
Db	431	tgttgaagatcagttcgtatgttaacccactcgttgacccaactgtgttcaagatcttcta	490		
OY	557	ctttcaacgaaggtttcttggtgtgagcaaaaacgaagaagcaaaatgacgca--aaagaag	614		
Db	491	ctttcaacgaaggtttcttggtgtgagcaaaaacgaagaagcaaaatgacgcaaaagga	550		
OY	615	aataagggcgacagcgaatgttgaatactatactcttcccttttcaataattatgaag	674		
Db	551	atnaacggcgacacaggaatgttgaaatcctaactcttcttcaataattatgaag	610		
OY	675	catctacgaagttattgtctcctcagaagcgatataattgaatgatttagaaaaataa	734		
Db	611	catctacgaagttattgtctcctcagaagcgatataattgaatgatttagaaaaataa	670		
OY	735	acaaatgaaggtgtcgcg-cgcacatctccgcgaagaagtgcacatcgtacgttcaagaacac	793		
Db	671	acaaatgaaggtgtcgcgcgcacacatcttcccgaaagtgccacactaaattgtacagcgctta	730		
OY	794	ttatlatcatgtacatt	809		
Db	731	attatttcggttaaaatt	746		
RESULT	4				
LOCUS	AL040542/c	767 bp	mRNA	EST	29-FEB-2000
DEFINITION	DKEP4341114.T1.434 (synonym: htes3)				Homo sapiens cDNA clone
ACCESSION	AL040542				
VERSION	AL040542.1	GI:5409490			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 767)
Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Koehler, et al.)
Unpublished (1999)
Contact: Koehler K

Am Klopferpitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ): Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charité,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No 5' sequence available.
This clone (DKFZp434i1114) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434i1114"
/clone_lib="434 (synonym: htes3)"
/issue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 219 a 158 c 168 g 221 t 1 others
ORIGIN

Query Match 65.0%; Score 650.6; DB 10; Length 767;
Best Local Similarity 99.4%; Pred. No. 4.6e-176;
Matches 653; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

131 ttgtgatgtcttcattcagtcggtcccaacagatcaaggcagatcatgatcccca 190
765 ttgtgatgtcttcattcagtcggtcccaacagatcaaggcagatcatgatcccca 706
191 tctgtgcaaaaaagcgttgaagtccttcggtcccgatcgtgtcagaagtaagttgg 250
705 tctgtgcaaaaaagcgttgaagtccttcggtcccgatcgtgtcagaagtaagttgg 646
251 ccgagtgatcactcagatggtatgagcagacatgacataatcttcttactgacat 310
645 ccgagtgatcactcagatggtatgagcagacatgacataatcttcttactgacat 586
311 ccgagtgatcactcagatggtatgagcagacatgacataatcttcttactgacat 370
585 ccgagtgatcactcagatggtatgagcagacatgacataatcttcttactgacat 526
371 tgcgagcagcagatggtcttcggtcccgatcagatgagatgagatgagatgagat 430
525 tgcgagcagcagatggtcttcggtcccgatcagatgagatgagatgagatgagat 466
431 gaacttaaaagtgctcaatctggaaaacgcttcggtcccgatcagatgagatgagat 490
465 gaacttaaaagtgctcaatctggaaaacgcttcggtcccgatcagatgagatgagat 406
491 taacggtgttaagatcagatggtatgagatgagatgagatgagatgagatgagat 550
405 taacggtgtgtgagatcagatggtatgagatgagatgagatgagatgagatgagat 346
551 ctttacttcacacagcgtttcgtgtgagcgaacaaacaggaagcgaacaaacaggaag 610
345 ctttacttcacacagcgtttcgtgtgagcgaacaaacaggaagcgaacaaacaggaag 286
611 agggaaataaagcgcagacggaacaaatgtgaaatacactacatcttcttcaatattt 670
285 agggaaataaagcgcagacggaacaaatgtgaaatacactacatcttcttcaatattt 226
671 gaagcattatcagaggtatattgtctcatgagcgagatacatatttgatattgaaaa 730

|||||
Db 225 GAAGCATTTATCAGGCTTATTGTCTCATACGCGATACATATTGTAATGATGTAAGAAA 166
QY 731 ataacaataaggggttcgcgcacatctcccgaaaagtgccacatgacgtctaa 787
Db 165 ATAAACAATAGGGGTTCGGCGACATTTCCCGAAAAGTCCACCTAATATTGTAAG 109

RESULT 5
AL598527/c 693 bp mRNA EST 14-AUG-2001
LOCUS DKFZp31301920_r1 313 (synonym: hlc2) Homo sapiens cDNA clone
DEFINITION DKFZp31301920 5', mRNA sequence.
ACCESSION AL598527
VERSION AL598527.1 GI:15161218
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann
S.
EST (Duesterhoeft, et al.)
Unpublished (1999)
Contact: Duesterhoeft A

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Am Klopferpitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ): Email: s.wiemann@dkfz-heidelberg.de;
sequenced by Olagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No 5' sequence available.
This clone (DKFZp31301920) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..693
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp31301920"
/clone_lib="313 (synonym: hlc2)"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiIb;
cDNA-collection"

BASE COUNT 173 a 167 c 156 g 194 t 3 others
ORIGIN

Query Match 63.7%; Score 638; DB 10; Length 693;
Best Local Similarity 99.7%; Pred. No. 1.9e-172;
Matches 638; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

362 aatagtgatgcgcgcagcagatggtcttcggtcccgatcagatgagatgagatgagat 421
Db 685 AATAGTGATGCGCGACCGAGTGTCTTCCCGCGCTCAATACGATTAATACCGCGC 626
422 caatagcagaaacttaaaagtgctcaatctggaaaacgcttcggtcccgatcagatgagat 481
Db 625 CACATAGCAGANCTTTAAAGTGCTCATCATTTGAAAACGTTCTTCCGGGCGCAAAACCT 566
482 caagatcttacgcgtgttgagatcagatggtatgagatgagatgagatgagatgagat 541
Db 565 CAAGATCTTACCGCGTGTGAGATCCAGTTGATGTAACCCACCTGTCACCACTGAT 506
542 cttaagcatttacttcacacagcgtttcgtgtgagcgaacaaacaggaagcgaacaaatg 601
Db 505 CTTGACATCTTTTACTTTCACCGCGTGTGAGATCCAGTTGATGTAACCCACCTGTCACCACTGAT 446
602 ccgcaaaaaaaggaataaagcgcagacggaacaaatgtgaaatacactacatcttcttcttc 661
|||||

Db 445 CCGCAAAAAAGGCAATAGGCGGACAGCAATGTTCACTACTCTCTCTTTTC 386
 QY 662 aatatttgaacattatcagaggtatgtctcatgagcggatataatttgaatgta 721
 Db 385 AATTATTTAGAGATTATTCAGGGTATTGTCTCATGAGCGATACATATTGAATGTA 326
 QY 722 tttagaataatacaaataggggttcgcgcacatttcccgcaaaagtgccacgtgag 781
 Db 325 TTGTGAAAAATTAACAATAGGGGTTCCGGCACATTTCCCGAATAAGTCCACCTGAGC 266
 QY 782 tctagaagaacattatcatatgacattacctaataaagaagcgtatagaagccct 841
 Db 265 TCTAGAAGAACATTATTATCATGACATTAACTTAATAAATAGCGGATACAGAGCCCT 206
 QY 842 ttctctcgcgcgttcgtctgagtagcgtgtaaaacctctgacacatgacgtcccgaga 901
 Db 205 TTGCTCTCGCGCGCTTTCGGTATGATGACGTGAAAACTCTGACACATGACCTCCGGAGA 146
 QY 902 cgtgcaacagctgtctctgtaagcgcagtcgcggagcagaacagcccgtaagggcgctcag 961
 Db 145 CGGTACACAGCTTTGCTGTAAAGCCGATGCCGGACAGACAAGCCGTCAGGGCGCTCAG 86
 QY 962 cgggtgttgccggtgtcgcggagcgtgcttaactatgctgcgc 1001
 Db 85 CGGGTGTGGCGGGGTGTCGGGGCTGCTTAACCTATGCGGC 46

RESULT 6
 AL042640/c 688 bp mRNA EST 29-PEB-2000
 LOCUS DKEZp434M0421_1 434 (synonym: htes3) Homo sapiens cDNA clone
 DEFINITION DKEZp434M0421 5', mRNA sequence.
 ACCESSION AL042640
 VERSION AL042640.1 GI:5422090
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 688)
 AUTHORS Blum H., Bauersachs, S., Mewes, H. W., Gassenhuber, J. and Wiemann, S.
 TITLE EST (Blum, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Blum H
 MIPS
 Am Klopferplatz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 Sequenced by LMU (Ludwig Maximilians University,
 Munich/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 No sl sequence available.
 This clone (DKEZp434M0421) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
 1..688
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKEZp434M0421"
 /clone_lib="434 (synonym: htes3)"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pSPort1; Site_1: NotI; Site_2: SalI"
 BASE COUNT 190 a 150 c 158 g 190 t
 ORIGIN

Query Match 63.6%; Score 636.6; DB 10; Length 688;
 Best Local Similarity 99.4%; Pred. No. 4.7e-172;
 Matches 639; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 145 ttcaagcttcggttcccaagatcaagggcgaagttaacatgatacccccattgtgtcaaaaa 204
 Db 688 TTTCAGCTCCGGTTCCCAAGATCAAGGGAGATTACATGATCCCATGTTGGCAAAAA 629
 QY 205 gcggttagctcttcggttcctcgatcggtgtgcaagaatgaagtgtgcccagtgatca 264
 Db 628 GGGGTAGCTCTCTCGGTCCTCGATCGTTGTGACAGAAAGTAAGTTGGCCGACATGTATCA 569
 QY 265 ctcaatgattatggcagcactgataattctcttaactgtcaatgccaatcgtaagtctt 324
 Db 568 CTCATGTTATGGACAGCACTGCATTAATCTCTTACATGTCATCCATCCGTAAAGATGCTTT 509
 QY 325 tctgtgactggtgagatgacccaacgaatcattctgagaatagtgtatgagggacgaagt 384
 Db 508 TCTGTACTGTGAGTACTCAACCAAGTATTTCTGAGANTAGTGTGCGGCGACCGAGT 449
 QY 385 tgccttcgcgcgcgtcaataacggaataacgcgcgcacatagagaacttaaaatg 444
 Db 448 TGCTCTGGCCGGCGTCATACGGGATATACCGGCCACATAGCAACTTTAAAGTG 389
 QY 445 ctcaatattggaataacggtcttcgcggagcaaacctcagaagatcttaccgtgttga 504
 Db 388 CTCATCATTTGAAAAAGCTTCTTCGGGCGCAAAACCTCTCAAGATCTTACCGCTTTGAGA 329
 QY 505 tccagttgagtgaacccaactgltgacccaactgactcttcagacatcttacttacc 564
 Db 328 TCCAGTTGATGTACCCACTGTCGCCACCACTGATCTTTCAGCATCTTTACTTTTCAC 269
 QY 565 agcgtcttcgtgtgagcaaaacagaagaagcaaaatgctgcgcaaaaaaaggaatgaaggcg 624
 Db 268 AGGTTTCTGCGTGACCAAAACAGAGAGCAAAATGCCGCAAAAAAGGAATTAAGCGCG 209
 QY 625 acaaggaatgttgaatactatactcttccttttcaatatattatgaagcattatcg 684
 Db 208 ACAAGCAAAATGTTGAATACATACATCTCTCTTCAATATTATGAAGCATTTATTCAG 149
 QY 685 ggtattgtctatgagcagatataattgaatgtaattgaagaataaacaataaggg 744
 Db 148 GGTATTGTGCTCTAGACGGATTCATATTGTAATGTAATTTAGAAAAATTAACCAATAGGG 89
 QY 745 gtccgcgcacattcccgaaagtgccacactgacgtctaaag 787
 Db 88 GTTCCGGGACATTTCCCGCAAAAGTGCACACTTAATTTGAAG 46

RESULT 7
 BE569261 914 bp mRNA EST 15-AUG-2000
 LOCUS 60131179F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3708859 5',
 DEFINITION mRNA sequence.
 ACCESSION BE569261
 VERSION BE569261.1 GI:9812981
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 914)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-f@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L146928 row: m column: 20
 High quality sequence stop: 685.

FEATURES

source
 location/Qualifiers
 1. .914
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3708859"
 /clone_1db="NCI-CGAP_Mam6"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 275 a 207 c 199 g 233 t

Query Match 60.7%; Score 607.2; DB 10; Length 914;
 Best Local Similarity 98.6%; Pred. No. 1.4e-163;

Matches 623; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Db 1 TCCGAACGATCAAGCGGAGTTACATGATCCCATGTTGCAAAAAAGCGTTAGCTCC 60
 QY 157 tcccaacgataaagcgagttacatgatacccaatgtgtgcaaaaagcggttagctcc 216
 Db 217 ttcggtccctccgattgtgtgcaagaatgtgtgcccgaagttatcactcatgtgtatg 276
 Db 61 TTGGGTCCTCCGATCGAGTACAGAAAGTGGCGCGAGTGTATCATCTCATGTTAAG 120
 QY 277 gcagcactgataattctcttactgtcatgcatccgtaagatgtcttctgtactgtg 336
 Db 121 GCAGCACTGATATTTCTTACTGTGATGCCATCCGTAAGATGCTTTCTGTGATGCT 180
 QY 337 gactactcaacgaatcatctgtgagaatgtgtgtgagcgagcagcgatgtcttgcgc 396
 Db 181 GAGTACTCAACCAAGTCAATCTGAGTAAGTATGCGCGAGCGAGTGTCTTGTGCCG 240
 QY 397 gctcaatcggagataatcccgccacatagagaactttaagttgctcatctatgta 456
 Db 241 GCCTCAATACGGATATATCCGCGCATAGAGAACTTTAAAGTGTCTCATATGGA 300
 QY 457 aaacgtcttcggggcgaaaactctcaagatcttccgctgtgtgagcagtcagatg 516
 Db 301 AAACGTTTTCGGGGCGGAAACCTCAAGATCTTACCGCTGTGAGATCCAGTTGAG 360
 QY 517 taacccactgtgaccccaactgcatcttcagcatcttcttacttcaacagcgttctg 576
 Db 361 TAAACCACTCGTGCACCCCAACTGATCTTCAGCATCTTTTCAACAGCGTTCTGG 420
 QY 577 tgaagcaaaaaggaaggaagaaatgcccgaaaaaggaaggaatagggcgacgaagaa 636
 Db 421 TGAAGCAAAAAGGAGGAAATGCGCAAAAAGGGAATAGGCGACACAGGAAATGT 480
 QY 637 tgaatacataactctcttcttcaatataatgaagcatttaccaggttatgtctc 696
 Db 481 TGAATCATCACTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGTTATGTCTC 540
 QY 697 atgagcgagatataattgaatgtatcttagaaaaaaataaagaatggggttcgcgcaca 756
 Db 541 ATGAGCGGATATATTTGAATGTATTAAGAAAAATAAACAATAGGGGTTCCGCGCACA 600
 QY 757 -tttcccgaaaagtcacccctgagcttaag 787
 Db 601 TTTTCCCGAAAAAGTGCCACTAATGTGAAG 632

RESULT 8
 ALO42540 702 bp mRNA EST 29-FEB-2000
 LOCUS ALO42540.1
 DEFINITION DKFZp341i1621.1 t1 434 (synonym: htes3) Homo sapiens cDNA clone
 DKFZp341i1621.1, mRNA sequence.

ACCESSION ALO42540
 VERSION ALO42540.1 GI:5421995
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 702)

AUTHORS Blum, H., Bauersachs, S., Mewes, H. W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Blum, et al.)

JOURNAL Unpublished (1999)

COMMENT Contact: Blum H

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

Sequenced by LMU (Ludwig Maximilians University, Munich/Germany) within the cDNA sequencing consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZp341i1621) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcententrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

location/Qualifiers
 1. .702

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DKFZp341i1621"

/clone_1db="434 (synonym: htes3)"

/tissue_type="testis"

/dev_stage="adult"

/lab_host="DH10B"

/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT

185 a 159 c 164 g 192 t 2 others

ORIGIN

Query Match 59.1%; Score 591.4; DB 10; Length 702;
 Best Local Similarity 96.2%; Pred. No. 4.7e-159;

Matches 637; Conservative 0; Mismatches 22; Indels 3; Gaps 3;

QY 129 gttgtatgtcttcaatcagctccggttcccaagcagtcagcgcg-aggatcatgaccc 187
 Db 701 GTATGGGTTCCATTCACAGCTCCCGGTTCCCAACGATCAAGCGAAGTTACATGATCC 642
 QY 188 ccatgtgtgcaaaaagcggttagctcttcggtccctccgcatc-gttgtcaagaagtaag 246
 Db 641 CCATGTTGTGCAAAAAGCGGTAGCTGCTGCTCCGATCGGTTGTCAGAAATGAAG 582
 QY 247 ttggccgagtgatatacatatgtgtatgtgacgacatgc-ataatcttctatctcat 305
 Db 581 TTGGCCGAGTGTATACATCATGATGTTATGACACACATGCAATTAATCTCTGTCAT 522
 QY 306 gccatccgttaagatgcttcttctgtgactgtgtgacttaacaaagtcattctgagaata 365
 Db 521 GCCATCCGTAAGATGCTTTTCTGTGACTGTGTGATCTCAACCAAGATCTTGAGATA 462
 QY 366 gtgtatgcygacgagagttgtcttgcgcgcggtcaataagggataataacgcgcacaca 425
 Db 461 GTGTATGCGGAGCGAGTGTGCTTGTCCCGGCTCAATACGGATTAATACCGCGCACA 402
 QY 426 tagaagaactttaaagtgctcatcatatgtgaaaaagttcttgcggcgaaaacttccaag 485
 Db 401 TAGCAAGACCTTTAAATAGTCTCATTTGAAAAACGTTCTTCGGGCGAAAACTCTCAAG 342
 QY 486 gatcttccgctgttgagatccagttcgatgtgaaacacacgcgtgcaccccaactgacatc 545
 Db 341 GATCTTACCGCTGTGAGATCCAGTTGATGTAACCCACTCGGCACCCCACTGATCTTTC 282
 QY 546 agcatcttcttacttaccagcgcttctgtgtgtagcaaaaaacgagaaggaatgtccgc 605
 Db 281 AGCATCTTTTACTTTCACGAGCGTTCTGGGTGAGCAAAAAAGCAAGGCAAAATGCGCG 222

OY	606	aaaaaagggaataaaggcgacacacgggaataatgttgaatactactactctctttcata	665
Db	221	AAAAAAGGAATAAGGGGACACGGAAATGTTGAATACTACTACTCTTTTCAATA	162
OY	666	ttatgaagcatcttccagggttattgtctcatagagcgatacalattgaatgatlla	725
Db	161	TTATGGAAGCATTTATTCAGGGGTTATTTGTCATWAGCGGATACATATTTGAATGATTTA	102
OY	726	gaaaaataacaataagggttcgcgcgcacatttccccgaaaatgccaacctgaagctla	785
Db	101	GAAAAATTAACCAAAATGAGGGGTTCCCGGACATATTTCCCGAAAAAGTCCACCTAAATGTGA	42
OY	786	ag 787	
Db	41	AG 40	
RESULT	9		
LOCUS	AG002183	C	
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: PQ624A20, genomic survey sequence.		
ACCESSION	AG002183		
VERSION	AG002183.1		
KEYWORDS	GI:2599226		
SOURCE	GSS.		
ORGANISM	Homo sapiens DNA, clone:PQ624A20.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 695)		
JOURNAL	Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y. Direct Submission Submitted (10-NOV-1997) to the DDBJ/EMBL/GenBank databases. Masahita Hattori, Kitasato University, Department of Science, JST Masahita Hattori, Kitasato University, Department of Science, JST (E-mail:hattori@ngc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561) 2 (bases 1 to 695) Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y. Homo sapiens genomic DNA, chromosome 21q Published only in database (1997) In press		
FEATURES	Location/Qualifiers		
SOURCE	1..695		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="21"		
	/clone="PQ624A20"		
	/map="21q"		
BASE COUNT	190 a 155 c 173 g 170 t 7 others		
ORIGIN			
Query Match	58.8%; Score 588.2; DB 13; Length 695;		
Best Local Similarity	95.1%; Pred. No. 3,9e-158;		
Matches 624; Conservative	0; Mismatches 30; Indels 2; Gaps		
OY	9	tttatccgctcatccagctcatctaatgttgcgcggaagctagaagtaagtgtcgcc	68
Db	694	TTTAAACCCNTCCATCCGTCGTTATTAATG-TTCCNGAGAACTAGATAGTATCCCC	636
OY	69	agttaatatttgcgaagaagtgttgccattgtctaaggcatcgttgttaacgctcgc	128
Db	635	AGTTAATAGTTTGGCAAGCGTTGTCACATGCTACAGGCAATGtGtGTCACGCTGTC	576
OY	129	gtttgtatagctctcatcagctccggttccccaagcataaggcgagttacatgaccc	188
Db	575	GTTTGGTAAGGCTTCAATTCAGCTCGGNTTCCCAAGAAACAGCGAGTATACATGAATGCC	516
OY	189	catgttgttcaaaaaagcggttagctcctctcgtgctccctcgatcgttgtcaagaagt	248
Db	515	CAATGTTGTCACAAAAGCGGTTTACGCTCTTCGCTCCGAGATGTTGTGAG-AGTAAGTT	457

OY	249	ggcgagatgttatcaactcaccagttaibgagacatgaattcttacttgcatgcg	308
Db	456	gccgcagctgtttatcacatgatgtttaibggcagcacatgcataatttcttactgcatgcc	397
OY	309	atccgtaagaagcttcttcctgtgactggtgcagtlactcaaccaagtcattctigaaatagtg	368
Db	396	atcccttagatgcttttcttgtatgcaggagactgacccaagatcattctgagatatg	337
OY	369	latgcgagcagcagatgtgctcttcgcccggcgatcaatacgggataataacggccacaatag	428
Db	336	tatggcgagccagcagatgttcttgcgccgctgcaatpacgggatatattacggccacatag	277
OY	429	cagaactttaaagtgccatcatcttggaanaacgittcttcggggcggaaacctcaagat	488
Db	276	cagaactttaaagtgccatcatcttggaanaacgittcttcggggcggaanaacctcaagat	217
OY	489	cttacgcctgttgatcatcagttgcagtgtlaaoccatcgtgtcacaccaatgatcttcagc	548
Db	216	ctttaccgctgttgatcatcagttgcagtgtlaaoccatcgtgtcacaccaatgatcttcagc	157
OY	549	alccttactcttaccagcgittcttggtgtgagcaaaaacaggaaggaatgccccgcaa	608
Db	156	atcttttacttcttaccagcgittcttggtgtgagcaaaaacaggaaggaatgccccgcaa	97
OY	609	aaagggaataaggcgagcacagcaggaatgttgatatctactcttccttttcaat	664
Db	96	aaagggaataaggcgagcacagcaggaatgttgatatctactcttccttttcaat	41
RESULT_10			
LOCUS	AG014918/c	658 bp DNA GSS	08-FEB-1999
DEFINITION	Homo sapiens genomic DNA, 21q region, clone: B335C24BB28, genomic survey sequence.		
ACCESSION	AG014918		
VERSION	AG014918.1 GI:3650162		
KEYWORDS	GSS.		
SOURCE	Homo sapiens DNA, clone:B335C24BB28.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-SEP-1998) to the DBUJ/EMBL/GenBank databases. Masahira Hattori, Kitasato University, Department of Science, Sagamihara 228, Japan (E-mail:hattori@ngc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561)		
REFERENCE	2 (bases 1 to 658)		
AUTHORS	Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.		
TITLE	Homo sapiens genomic DNA, chromosome 21q		
JOURNAL	Published Only in Database (1998) In press		
FEATURES	Location/Qualifiers		
source	1..658 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="21" /clone="B335C24BB28" /map="21q"		
BASE COUNT	164 a 150 c 158 g 183 t 3 others		
ORIGIN			
Query Match	57.6%; Score 576.8; DB 13; Length 658;		
Best Local Similarity	96.5%; Pred. No. 7.3e-155;		
Matches 610; Conservative 0; Mismatches 19; Indels 3; Gaps			
OY	328	gtgactgtgtgactactcaacaaatcattctgagaatgtgtatgcgcgacaggattgc	387
Db	658	gtgactgtgtgactactcaacaaatcattctgagaatgtgtatgcgcgacaggattgc	599
OY	388	tcttgcgcggcgatcaatagcgataatattacggccacatagacagaatttaaagtgtcc	447

```

|||||
Db 598 TCCTGGCCCGGCGTCAATACGGGATTAATACCGCCGACATGACGAACTTT-AAAGTCTC 540
Qy 448 atcatggaagaacgtcttcctggggcgaaacatctcaagatctaccgc--tgtagat 505
|||||
Db 539 ATCAATTTGAAAACGTTCTTCTGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTTGAAT 480
Qy 506 ccaattcagatgaaccactcgtgcacccaactgactctcagcatcttacttaccaca 565
|||||
Db 479 CCAATTCGATGTAACCACTCGTGCAACCACTGATCTTACGATCTTACTTTCACCA 420
Qy 566 ggcgttcctgggtgagcaaaacaggaagcacaatgcccgaacaaagggaataagggcga 625
|||||
Db 419 GCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCCAAAAAGGAATTAAGGCGCA 360
Qy 626 caggaagaatgtgaatcactcactcttcttcaatatattatgaacattatcag 685
|||||
Db 359 CACGGAATGTTGAATACATACCTCTTCCCTTTTCAATTAATTAACCATTTATACAG 300
Qy 686 gtatctgctcagtagcgagatacatacttgaaatgtaattagaacaaataaataaggg 745
|||||
Db 299 GTATTGTCTCATGACGAGATACATATTGATGATTTAGAAAAATAACAATTAAGGG 240
Qy 746 ttccgcgcacattcccgaaagtgccacctgacgctcctaagaacattatcatga 805
|||||
Db 239 TTCGGGCACTTTCCCGAAAAAGTGCACCTGACGCTTAAGAACCATTTATATCATGA 180
Qy 806 cattaaccataaataagagcgtatcagagggcccttcgtctgcgcttcgtagatg 865
|||||
Db 179 CATTAACCTATAAATAAATGAGCGATACAGAGGCCCTTTCGTCTGCGCTTTCGGTGAAG 120
Qy 866 acggtgaagaacctctgacacatgacgtcccgagagacggtcacaactgtgctgaagcg 925
|||||
Db 119 ACGGTGAATAACCTCTGACACATGACAGCTCCCGAGAGGATCACACGTTCTCTGTAAGCG 60
Qy 926 atgcggagagcagacaagccgctcagggcg 957
|||||
Db 59 ATGCCGGAGAGCAAGCAAGCCGCTACCTCCG 28

```

RESULT 11
AG000364/c 793 bp DNA GSS 06-FEB-1999
LOCUS Homo sapiens genomic DNA, 21q region, clone: P0119B8X30, genomic
DEFINITION survey sequence.
ACCESSION AG000364
VERSION AG000364.1 GI:2579172
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: P0119B8X30.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-1997) to the DDBJ/EMBL/Genbank databases.
Masahita Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail: hattori@jgc.ims.u-tokyo.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
2 (bases 1 to 793)
REFERENCE AUTHORS Hattori, M., Ishii, K., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1997) In press
FEATURES Location/Qualifiers
1..793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="P0119B8X30"
/map="21q"
BASE COUNT 205 a 185 c 207 g 180 t 16 others
ORIGIN

```

Query Match 55.9%; Score 559.2; DB 13; Length 793;
Best Local Similarity 97.7%; Pred. No. 8,6e-150;
Matches 594; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

```

Qy 1 cctgaacttattccgcctcatccagctcattatgttgcgggaagtagagtaag 60
|||||
Db 634 CCGTAACCTTTATCCGCCCTCCATCAGCTATTAATTTCTCCGGGAAGCTAGTAAGT 575
Qy 61 agtgcagatgaatagtttgcgaacgttggtgcattgacaaagc-cacgtgtgtc 119
|||||
Db 574 AGTTGCCAGTAATAGTTTGGCGCAACGTTGTGCATTCCTACAGGCAACGCGTGTG 515
Qy 120 acgtcgtcgtttggtatagctcattcagctccggttcccaagatcaagggagttac 179
|||||
Db 514 ACGCTGTCGTTTGGTATAGGCTTATGCTTCGCGTTCCCAAGATCAAGGCGAGTTAC 455
Qy 180 atgattccccaatgtgtgcaaaaagcggttagctccttcgtgctccgattcgttgcag 239
|||||
Db 454 ATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTTCGCTTCGATCGTTGTCAG 395
Qy 240 aagtaagttgcgcgagtgatcatcactcattgtatgtgacagcactgcataatctctac 299
|||||
Db 394 AAGTAAGTTGGCCGAGGTATACATCATAGTATGAGACACTGCATTAATCTCTTAC 335
Qy 300 tgtatgcacatcgcgaagatgtcttctgtgactgtgtgactcacaacgaatcattctg 359
|||||
Db 334 TGTCATGCAATCCGTAAAGATCTTTCTGTGACGTGGAGTACCAACCAAGTATCTGTG 275
Qy 360 agaataagtgatgctgcgcgcagcagctccttcgcgcgcgcacaaataggaataacgc 419
|||||
Db 274 AGAATAGAGTATGCGCGACGACGAGTGTCTTCCCGGCTCAATACGGATTAATACCG 215
Qy 420 gccacatagcagaacttaaaagtgcatacatgtgaaaa-cgttcttcggggcgaaac 478
|||||
Db 214 GCCACATAGCAGAACTTTAAAGTGTCTTCATGTGAANNCTGTTCTTCGCGGCAAAAC 155
Qy 479 t-cctcaagatctaccgcgttgtagatccagttcgaatgaacccactgtgacccaac 537
|||||
Db 154 TNCCTAAGGATCTTACCGCTGNTGAGATCCAGTTCGATTAACCCACTCTGTGACCCAAC 95
Qy 538 tgatctcagcatcttacttacttaccagcgttctcgtgtgaggaacaaaggaagcga 597
|||||
Db 94 TGATCTTCAGATCTTTACTTTCACCAAGGCTTTCNCGGTGNCMAAAGCAAGAGCA 35
Qy 598 atgcgcgc 605
|||||
Db 34 AATGCCCTC 27

RESULT 12
AG009765/c 696 bp DNA GSS 14-APR-1999
LOCUS Homo sapiens genomic DNA, 21q region, clone: P864SPN15, genomic
DEFINITION survey sequence.
ACCESSION AG009765
VERSION AG009765.1 GI:3289751
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: P864SPN15.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1998) to the DDBJ/EMBL/Genbank databases.
Masahita Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan

(E-mail: hatori@ngc.ihs.u-tokyo.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
On Feb 5, 1999 this sequence version replaced g1:2706702.
AG003576: Submitted (19-Dec-1997).

FEATURES

Location/Qualifiers

1.696
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="P8G4SPN15"
/map="21q"
BASE COUNT 183 a 165 c 185 g 162 t 1 others
ORIGIN

Query Match 55.3%; Score 554; DB 13; Length 696;
Best Local Similarity 99.5%; Pred. No. 2,6e-148;
Matches 587; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1 cctgcaacttaccgctccatccatcagtcattatattgtg-cgggaagctagaag 59
DB 620 CCTGCACTTATCCGCTCCATCCATCCTATTAATGTGCGGAGCACTAGAGTAAG 561
QY 60 taatgcccgaatgaatgattgagcaacgttgtgcatgctaacag-catcggtg 117
DB 560 TAGTTCCGCAATTAAATGTTGGCCAAAGTGTGGCCATGTACAGGCCATCGTG 501
QY 118 tcacgctcgttgtgtgattgattcattcagtcgcgcgtcccaacatcaagcgag 177
DB 500 TCAGGCTCGTGTGGTATGGCTTCATTCAGCTCCGCGTCCCAAGCATCAAGCGAG 441
QY 178 acatgatcccatgtgtgtgcaaaaagcggttagctccttcgctccgcatcgtgt 237
DB 440 ACATGATCCCATGTTGTGCAAAAAGCGGTAGCTCTCGGCTCCGATCGTGTGC 381
QY 238 agaagtaagtggccgagtgattcactaatgatttgagcagcagcgaattctct 297
DB 380 AGAAGTAAGTGGCGGAGTGTATCACTCATGATGAGAGCATGCAATATCTCTT 321
QY 298 actgtcatccatccgaatgaatgtcttctgtgactgtgagtaactaacgaatc 357
DB 320 ACTGTCAATGCCATCCGTAAGATGCTTTCTGTGACTGTGATGATCAACCAAGTATTC 261
QY 358 tgaagaatagtgatgagcgagcagctgtcctgtcccgcgatcaataagataacc 417
DB 260 TGAGATATGTATGCGGCGACGAGTGTGCTTCCGCGGCTCAATAGGATATATACC 201
QY 418 ggcgcacatagcaaacatttaaaagtctcatcttggaagaaacgttctggggcg 477
DB 200 GCGCCACTACCAAACTTTAAAGTCTCATATGGAACCGTTCTTGCGGGCGAAAA 141
QY 478 ctctcaagaatcctaccgctgttagatccagttcgatgttaaccacatcgcccaac 537
DB 140 CTCTCAAGGATCTTACCCGCTTTAGATCCAGTTGATGTAACCATCTCGGCGACCAAC 81
QY 538 tgatctcaagcatcttacttaccagcagcttctgggtgagcaaaaac 587
DB 80 TGATCTTACGATCTTTTACTTTCACAGCGCTTCTGGGTGAGCAAAAAC 31

RESULT 13

LOCUS BE573230 696 bp mRNA EST 15-AUG-2000
DEFINITION 601333128F2_NCL_CGAP_Mam6 Mus musculus cdna clone IMAGE:3710479 5',
mRNA sequence.
ACCESSION BE573230
VERSION BE573230.1 GI:9816950
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 696)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM8933 row: a column: 08
High quality sequence stop: 696.

FEATURES

Location/Qualifiers
1.696
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3710479"
/clone_lib="NCL_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; Salt:
Site_2: NotI; Cloned unidirectionally; Primer: Oligo dT.
Library constructed by Life Technologies, Inc. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 196 a 157 c 152 g 191 t
ORIGIN

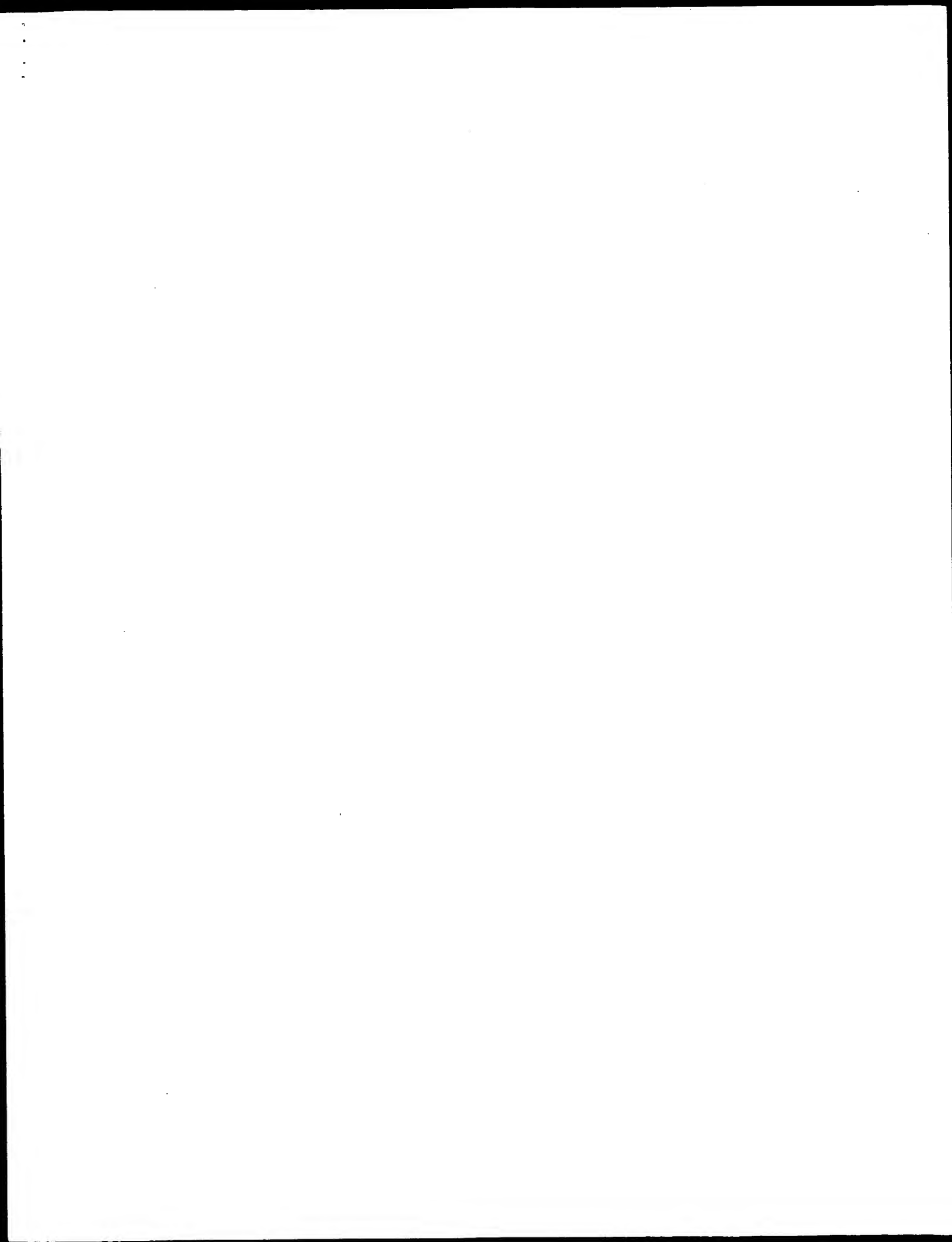
Query Match 54.8%; Score 548.4; DB 10; Length 696;
Best Local Similarity 98.9%; Pred. No. 1.1e-146;
Matches 615; Conservative 0; Mismatches 1; Indels 6; Gaps 6;

QY 151 tccggttcccaagaatcaagcgagttacatgatccccatgttgtgcaaaaacggt 210
DB 1 TCCGGTTCCTCCAAAGATCAAGCGAGTGTACATGATGCCCATG-TGTGCAAAAAGCGGTT 59
QY 211 agctccttgagctccgactcgtgtgtgtcagaagtaagtgtgcccgaatgtatcaact 270
DB 60 AGCTCTTCCGCTCTCGATCGTGTGCAAGTAAG-TGCCGCACTGTATCACATCGATG 118
QY 271 gttatgcaagcactgataatctcttactgtcatgcatccgtaagatgtcttctgt 330
DB 119 GTATGCGACACTGCATATATCTTACTGTCATGTCATCCGTAAGATGCTTCTGTG 178
QY 331 actgtgtgactcaacacgaatcattctgagaatgtgtatgagcgagcaggtgtct 390
DB 179 ACTGTGTGACTCAACCAAGTCTTGTGAAATGTATGATGGGCGAGCTGTCTT 238
QY 391 tggccgagcgtcaataaggaataataccgcgcacatagcagaacttaaaatgtctalc 450
DB 239 TGCCTCCGCTCAATAGGGATTAATACCGCCCATAGCAAACTTTAAAGTCCGATC 298
QY 451 attgaaaacgttcttcggggcgaaaactctcaagatcttaccgctgttgagatccagt 510
DB 299 A-TGGAAAAGCTTCTCGGGGCGAAAACTCTCAAGATCTTACCGCTGTGAGATCACT 357
QY 511 tcatgttaaccacactcgtgacccaactgatcttcaagcatcttcaaccagcgtt 570
DB 358 TCGATGTAAACCCACTGCTGACCACTGATCTTTCAGCATCTTTTACTTTCACACACG 417
QY 571 tctgggtgagcaaaaacgagaagcaaaatgcccgaaaaaggaataagggcgacaag 630
DB 418 TCTGGGTGACCAAAAGGAAAGGCAAAATGCGCAAAAGGAAATAGGGGCGACACGG 477
QY 631 aatgttgaatactatcactcttcttcaatattatgaagcaattatcagggttat 690
DB 478 AATGTGAAATACTCATACCTCTTCTTTCATATTA-TGAAGCATTTATCAGGGTA- 535


```
/clone_lib="434 (synonym: htesj)"  
/tissue_type="testis"  
/dev_stage="adult"  
/lab_host="DH10B"  
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"  
BASE COUNT      152 a      123 c      124 g      155 t  
ORIGIN
```

```
Query Match      54.5%; Score 545.6; DB 10; Length 554;  
Best Local Similarity 99.3%; Pred. No. 6.5e-146;  
Matches 548; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 226 ccgacgtgtgcaagaagttgagccgcaagtgtatcatcactcaatggtatgcaacactg 285  
    |||||||  
DB 554 CCGATCGTTGTCAGAGTAAGTGGCCGACGTTATCATCATGTTATGCGACACTG 495  
    |||||||  
QY 286 cataattcttactgtcatccatccgtaagatgctttctgtgactgtgagtaactca 345  
    |||||||  
DB 494 CATAATCTTACTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCA 435  
    |||||||  
QY 346 accaagtcattctgaagaatagtatgtgagcgacgagttgctcttgcccggtcaata 405  
    |||||||  
DB 434 ACCAAGTCTTCTGAGATAGTATGCGGACCGAGTTGCTTGGCCGCGTCAATA 375  
    |||||||  
QY 406 cgggataataccgcgcacatagcagaacttaaaagtgtcatcatgtgaaacgtttct 465  
    |||||||  
DB 374 CGGATATATTCGGCGCCATAGCAGAACTTTAAAGTGTCTATCATTTGGAAAAGTTCT 315  
    |||||||  
QY 466 tcggggcgaaaaactctcaagatcttaccgctgttgagatccagttcgaltglaaccact 525  
    |||||||  
DB 314 TCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCACT 255  
    |||||||  
QY 526 cgtgacccaactgactcttcgacatcttttactcttcaccagcgtttctggtgagcaaaa 585  
    |||||||  
DB 254 CGTGACCCAACTGATCTTCAGACATCTTTTACTTTCACACAGCGTTCTGGGTGAGCAAAA 195  
    |||||||  
QY 586 acaggaagcacaatgcccgaacaaaaggaataagggcgacagcaaatgttgaatactc 645  
    |||||||  
DB 194 ACAGGAAGCGCAAAATGCCGCAAAAAGGGAATAGGGCGACACGAAATGTGATACTC 135  
    |||||||  
QY 646 ataactctctcttccaatatatttgaagcaattatcaagggtatgtctcatgagcgga 705  
    |||||||  
DB 134 ATACTCTTCTTTTCAATATTTATTGAAAGCATGTGATGAGGTTATGTCTCATGAGCGGA 75  
    |||||||  
QY 706 tacatattgagtatttgaagaaataaacaataaggggtccgcgacatttcccgga 765  
    |||||||  
DB 74 TACATATTGATATTTAGAAAAATACCAATAGGGGTTCCGCGCACATTTCCGCCGA 15  
    |||||||  
QY 766 aaagtgcacact 777  
    |||||||  
DB 14 AAGTGCACACT 3
```

Search completed: January 17, 2002, 14:58:45
Job time: 8646 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:25:58 ; Search time 5101.32 Seconds
(without alignments)
3233.904 Million cell updates/sec

Title: us-09-740-211-14_COPY_1_1000

Perfect score: 1000
Sequence: 1 cgcacctgcagcagctgcg.....cagccctcattggagcccta 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824858755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pt: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rnd: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rnd: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	4999	6	AR138378
2	1000	100.0	4999	6	AR146888
3	736.2	73.6	11933	6	AR138377
4	736.2	73.6	11933	6	AR146887
5	592.2	59.2	5035	6	AR034084
6	592.2	59.2	5035	6	AR071306
7	592.2	59.2	5035	6	AR112722
8	590.6	59.1	8831	6	E00527
9	590.6	59.1	8967	6	AX052730
10	590.6	59.1	8967	6	171409
11	590.6	59.1	8967	9	HSEVITIR
12	590.6	59.1	9029	9	HUMFVITIC
13	589	58.9	9009	6	AR003710
14	589	58.9	9009	6	AR029065
15	589	58.9	9009	6	AR126884
16	589	58.9	9009	6	I31901
17	589	58.9	9009	6	I63424
18	589	58.9	9009	9	HUMFVITIC
19	586	58.6	8241	6	A05328
20	586	58.6	8241	6	A07042
21	585.8	58.6	7440	6	I08345
22	585	58.5	7272	6	I05404
23	583.4	58.3	7056	6	I27063
24	579.2	57.9	9354	6	AR003585
25	577.8	57.8	7440	6	I08457
26	528	52.8	4278	6	I08644
27	528	52.8	4281	6	I08643
28	528	52.8	4548	6	I08642
29	528	52.8	4551	6	I08641
30	528	52.8	4670	6	AR110040
31	528	52.8	6999	6	I77105
32	513.8	51.4	1993	6	I02047
33	454.2	45.4	7032	4	AF016234
34	452.6	45.3	7145	4	AF049489
35	451.4	45.1	7493	6	AR003712
36	451.4	45.1	7493	6	AR029067
37	451.4	45.1	7493	6	AR126886
38	451.4	45.1	7493	6	I63427
39	451.4	45.1	7493	10	M0SCVITII
40	424.8	42.5	6539	4	SS049517
41	423.8	42.4	4334	6	AR029098
42	423.8	42.4	4334	6	AR126917
43	423.8	42.4	6402	6	AR029097
44	423.8	42.4	6402	6	AR126916
45	380.2	38.0	553	4	AF180523

ALIGNMENTS

RESULT 1
AR138378
LOCUS AR138378 4999 bp DNA
DEFINITION Sequence 14 from patent US 6200560.
ACCESSION AR138378
VERSION AR138378.1 GI:14480723
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 4999)
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated virus vectors for expression of factor VIII by target cells
JOURNAL Patent: US 6200560-A 14 13-MAR-2001;
FEATURES
source location/Qualifiers
1..4999
/organism="unknown"

BASE COUNT 1375 a 1151 c 1155 g 1318 t
ORIGIN

[illegible][illegible]


```

|||||
Db 1 CACCTCGGGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAACCCGGGCGTGGGGGAGAC 60
QY 73 ttgtgtcgcgcgcgcctcagtgagcgagcgagcgagggagggagtgagcgaacccatc 132
Db 61 tttgtgtcgcgcgcgcctcagtgagcgagcgagcgagggagggagtgagcgaacccatc 120
QY 133 actaaggagtgctcctcgccgcgcgcgcgtggtggtcgcgcggtgaa-----ctcggaagt 185
Db 121 ACTAGAGGTTCTCTGGCGCGCGCGAGGAGATGTTGTTCTTAATACCATCCAGGGAATGT 180
QY 186 gatgtgtgtactgtctc-----gcttttcccgaggtgtggggagaaacgtatataagt 242
Db 181 TGTGTTCTTAATACCATCCAGGGAATGTTGTTCTTAATACCATCTACAGTTATGTGTT 240
QY 243 gcagatgcgcgcgcgcctccttcttgcgaacgggt-----tgcgcgcgcg 290
Db 241 AAGAAGATATTATAGAGCAGTCTTTCTGCACACAGATACCTTTCCGGGTCCGCCCT 300
QY 291 cgcgcagtgagtgagcgaggaatgtgttcttaataaccatgcctcaggaagtgttctgt 350
Db 301 AGGCAAGTAAGTGGCGTGTGTGTTCCCGCGGCGCTTTACGGGTATGCGCCT 360
QY 351 tcttaataccatctactgacacgtacatcaacttcttcttcttccacaggtatcga- 409
Db 361 TCGGTGCTTGAATTAATCACTGACATCCACTTTTCTTTCTTCCACAGGTATGAT 420
QY 410 tccacatgcaataagagatcctcaccctgtcttcttctgtgcttcttgcgttcttctt 469
Db 421 TCCACCATCAATATAGCTTCCACCTGCTCTTCTGTGTGCTTTTCGATTCGCTT 480
QY 470 agtgccaccagaagatactactctgtgtgagtgagacgtgacaggaatataagaaagt 529
Db 481 AGTGCCACCAGAAAGATTAATCACTGGGTGAGAGTGAATGATGAGGATATATGCAAGT 540
QY 530 gatttcggtgagctgcgtgtgtgagcgaagaattcctcctctagaagtgccaaatcttcca 589
Db 541 GATCTCGGTGAGCTGCTGCTGTGAGCCAAAGATTTCTCTAGAGTCCAAATCTTTTCCA 600
QY 590 tccaacacctgcgtgtgtacaaaagaactgtgtgtgagaaatagaagatcaacctttc 649
Db 601 TTCACACCTCAGTGTGTGTACAAAAGACTCGTTGTGATGATTCACGGATCACCTTTTC 660
QY 650 aacatgcgttaagcgaagccacctgtgagtggtctgtcgtaggttctaccatccaggtcag 709
Db 661 AACATCGCTAAGCCAAAGCCACCTGTGATGGGTCTGATGAGTCCATCCATCCAGGCTGAG 720
QY 710 gttatgatacagtggttaataaactaagaacatggttccatccctgtgaagcttcat 769
Db 721 GTTTATGATACGTGTGTATTAACCTTAAGAACATGAGCTTCCCATCCCTGATGCTTCAT 780
QY 770 gctgtgtgtatcctacttggaagactctcagagagctgaatagatgacagacagt 829
Db 781 GCTGTGGTGTATCTCTACTGGAAGCTTCTGAGGAGCTGATATGATGATGATGATGAT 840
QY 830 caaagggagaaagaatagataaagcttccctgtgtgagcgaatataatgctgtgcag 889
Db 841 CAAAAGGAGAAAGAGATGATTAAGTCTTCCCTGTGGAACCATACATATGATGTGGCAG 900
QY 890 gtccgaaagaagaatggtcgaatgacctcagccacgtgtcctactatataatcct 949
Db 901 GTCTGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 950 tctcatgtgacgtgtglaaagaacttgaatcaggtccatctatgagcccta 1000
Db 961 TCTCATGTGAGACTGTGTAAGACTTGAATTCAGGCTCATTCATGAGGCCCTA 1011

```

```

VERSION AR034084.1 GI:5949689
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Vooberg, J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;
FEATURES
Source 1..5035
BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN
Query Match 59.2%; Score 592.2; DB 6; Length 5035;
Best Local Similarity 97.9%; Pred. No. 8..9e-155;
Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 388 tctttctccacagratatccaccatgcaaatgagcttccaccgtcttcttct 447
Db 7 TCGAGTTGAACATTTGTAAGAGCCACCATGGAATAGACTCTCCACGCTTTCTTCT 66
QY 448 gtgccttctgcatctcgtcttaagtgcaccagaagaatactactggtgagtgagact 507
Db 67 GTGCTTTTGGGATTTGCTTGTAGTGCCACAGAGATTAATCTGCTGCTGATGGAAGT 126
QY 508 gtcatggaactatatacgaagtgatctcgtgtgagctgtcgtgtgagcgaagatctcc 567
Db 127 GTCATGGGACATATGCAAGTGTATCGGTGAGCTGCTGTGAGCAGCAAGATTTCCTCC 186
QY 568 tagagtgccaaatcttcttcatcaaacctcagtcgtgttacaataaagatctgttgt 627
Db 187 TAGAGTCCAAATCTTTCCATTTCAACACCTCAAGTGTGTATGATGATGATGATGAT 246
QY 628 agaattcaggaatcccttccaacatcgtctaaagcgaagccacctgagtgagtgctgt 687
Db 247 AGAATTCAGGATCACCTTTCAACATCGCTAAGCCAGCCACCTGGATGGGTGCTGCT 306
QY 688 aggtccatcaatccagcgtgaggttatgatatacagtggttcaataaataagaatgag 747
Db 307 AGGTCTTACCATCCAGCTGAGGTTATGATGATGATGATGATGATGATGATGATGAT 366
QY 748 ttccacatccgtcagcttcatgctgtgtgtatcctacttggaagctctcgaaggagc 807
Db 367 TTCCATCCTGTGATGCTTCTATGCTGTGTATCTTACTGGAAGGCTTTGAGGAGAC 426
QY 808 tgaatatgatatacagaaccagtcgaagggagaaagaatgataaagcttccctgtgtg 867
Db 427 TGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 486
QY 868 aagccatataatgctgtcagagtgctgaaggaatggttcaatgctcctgagccact 927
Db 487 AAGCATACATATGCTGAGGAGTCTGAAAGAGATGATGATGATGATGATGATGATGAT 546
QY 928 gtgccttaccatcatatcttctcatgtgagcctgtglaaagaacttgaatcagacct 987
Db 547 GTGCTTACTTCAATATCTTCTCATGTGAGACCTGTAAGAAAGCTTGAATTCAGGCT 606
QY 988 catltagagcccta 1000
Db 607 CATTTGAGGCCCTA 619

```

RESULT 5
 LOCUS AR034084 5035 bp DNA
 DEFINITION Sequence 1 from patent US 5869292.
 ACCESSION AR034084

RESULT 6
 LOCUS AR071306 5035 bp DNA
 DEFINITION Sequence 1 from patent US 5910481.
 ACCESSION AR071306
 VERSION AR071306.1 GI:7222194
 KEYWORDS Unknown.

ORGANISM Unknown
Unclassified.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Voorberg,J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5910481-A 1 08-JUN-1999;
FEATURES Location/Qualifiers
source 1..5035
BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN

Query Match 59.2%; Score 592.2; DB 6; Length 5035;
Best Local Similarity 97.9%; Pred. No. 8.9e-155;
Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 388 tcttttcccaagatcgcattccaccatgacaatagagctccaccctgcttctt 447
DB 7 TCCAGTTGACATTGTGTAGCAAGCCACCAGTAAGATAGAGCTCCACCTGCTTCTTCT 66
QY 448 gtgcctttgcattctgctttagtgcaccagaagaataactcctgggtgcagtgaact 507
DB 67 GTGCTTTTGGATTGCTTGTAGTGCACACAGAAATACCTGCGTGACAGTGAAC 126
QY 508 gtcatggactatatagcaagtgatctcgttgagctcctggtgagcaagatttctctcc 567
DB 127 GTCATGGGACTATATGCAAAAGTATCTGGTGAGCTCCTGTGAGCAAGATTTCTCC 186
QY 568 tagagtgcaaaatctttccatcaaacctcagctcgtgtcaaaaagaactcgttgt 627
DB 187 TAGAGTGCACAAATCTTTTCCATTCACACCTCAGTGTGACAAAAGACTGTTGT 246
QY 628 agaattcagagatcaacttttcaacatcgttaagcaagaccctgagtgctgtc 687
DB 247 AGAATTCACGGATCACCTTTTCAACATCGCTAACGCCAACGCCCTGATGGGTCTGT 306
QY 688 aggtccatccatccagagctgagttttagatagagtgatcacttaagaacatggc 747
DB 307 AGGTCTTACCATCCAGGCTGAGTTTATGATACAGTGTGCTTACCTTAAAGAACATGGC 366
QY 748 ttccatcctgtcagctcctcgtgtgtgtatctcactgaaagcttctgagggagc 807
DB 367 TTCCCATCTCTCAGTGTCTTATGCTGTGTATCTCCTACGGAAGCTCTTGAGGAGC 426
QY 808 tgaatatgatatcagaccagtcgaagggagaaagaatgataaagtccttccctgtg 867
DB 427 TGAATATGATGATCAGACCACTCAAGGGAGAAAGAGATGATAAGTCTTCCCTGTGG 486
QY 868 aagccatcatatgtctgagcagtcgtcgaagaagaatggtccaaatgctccttgaacct 927
DB 487 AAGCCATACATATGTCTGGCAGGTCTGAAAGAGATGCTCAATGGCTCTTGACCCACT 546
QY 928 gtgccttaccatcatatcttctcattgagcctgtaaaagactgaaatcaagcct 987
DB 547 GTGCCTTACCTACATCATATCTTCTCATGTGACCTGTTAAAGACTTGAATTGAGCCT 606
QY 988 cattgagacccta 1000
DB 607 CATTGAGACCCCTA 619

RESULT 7
AR112722 5035 bp DNA PAT 16-MAY-2001
LOCUS AR112722 5035 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6130203.
ACCESSION AR112722
VERSION AR112722.1 GI:14092622
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
TITLE Unclassified.

AUTHORS Voorberg,J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 6130203-A 1 10-OCT-2000;
FEATURES Location/Qualifiers
source 1..5035
BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN

Query Match 59.2%; Score 592.2; DB 6; Length 5035;
Best Local Similarity 97.9%; Pred. No. 8.9e-155;
Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 388 tcttttcccaagatcgcattccaccatgacaatagagctccaccctgcttctt 447
DB 7 TCCAGTTGACATTGTGTAGCAAGCCACCAGTAAGATAGAGCTCCACCTGCTTCTTCT 66
QY 448 gtgcctttgcattctgctttagtgcaccagaagaataactcctgggtgcagtgaact 507
DB 67 GTGCTTTTGGATTGCTTGTAGTGCACACAGAAATACCTGCGTGACAGTGAAC 126
QY 508 gtcatggactatatagcaagtgatctcgttgagctcctggtgagcaagatttctctcc 567
DB 127 GTCATGGGACTATATGCAAAAGTATCTGGTGAGCTCCTGTGAGCAAGATTTCTCC 186
QY 568 tagagtgcaaaatctttccatcaaacctcagctcgtgtcaaaaagaactcgttgt 627
DB 187 TAGAGTGCACAAATCTTTTCCATTCACACCTCAGTGTGACAAAAGACTGTTGT 246
QY 628 agaattcagagatcaacttttcaacatcgttaagcaagaccctgagtgctgtc 687
DB 247 AGAATTCACGGATCACCTTTTCAACATCGCTAACGCCAACGCCCTGATGGGTCTGT 306
QY 688 aggtccatccatccagagctgagttttagatagagtgatcacttaagaacatggc 747
DB 307 AGGTCTTACCATCCAGGCTGAGTTTATGATACAGTGTGCTTACCTTAAAGAACATGGC 366
QY 748 ttccatcctgtcagctcctcgtgtgtgtatctcactgaaagcttctgagggagc 807
DB 367 TTCCCATCTCTCAGTGTCTTATGCTGTGTATCTCCTACGGAAGCTCTTGAGGAGC 426
QY 808 tgaatatgatatcagaccagtcgaagggagaaagaatgataaagtccttccctgtg 867
DB 427 TGAATATGATGATCAGACCACTCAAGGGAGAAAGATGATAAGTCTTCCCTGTGG 486
QY 868 aagccatcatatgtctgagcagtcgtcgaagaagaatggtccaaatgctccttgaacct 927
DB 487 AAGCCATACATATGTCTGGCAGGTCTGAAAGAGATGCTCAATGGCTCTTGACCCACT 546
QY 928 gtgccttaccatcatatcttctcattgagcctgtaaaagactgaaatcaagcct 987
DB 547 GTGCCTTACCTACATCATATCTTCTCATGTGACCTGTTAAAGACTTGAATTGAGCCT 606
QY 988 cattgagacccta 1000
DB 607 CATTGAGACCCCTA 619

RESULT 8
E00527 8831 bp RNA PAT 29-SEP-1997
LOCUS E00527 8831 bp RNA PAT 29-SEP-1997
DEFINITION Complete cDNA encoding human factor VIII.
ACCESSION E00527
VERSION E00527.1 GI:2168806
KEYWORDS JP 1985243023-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 8831)
AUTHORS Danilov,J.K., Richiyoado,M.R., Goodson,A.B. and Urtiama,A.U.
TITLE MANUFACTURE OF FUNCTIONAL VIII FACTOR

Keyl, B., Seeburg, P.H., Smith, D.H., Hollingshead, P., Wion, K.,
 Delwart, E., Tuddenham, E.G.D., Vohar, G.A. and Lawn, R.M.
 Expression of active human factor VIII from recombinant DNA clones
 JOURNAL Nature 312 (5992), 330-337 (1984)
 MEDLINE 85061548

COMMENT Data kindly reviewed (20-MAR-1986) by W. Wood.
 FEATURES Location/Qualifiers

source

1..8967

/organism="Homo sapiens"

misc_feature

/db_xref="taxon:9606"

sig_peptide

/note="5' untranslated region"

CDS

/product="signal peptide (aa -19 to -1)"

/note="factor VIII precursor"

/codon_start=1

/protein_id="CAA25619.1"

/db_xref="GI:31499"

/translation="MOELSTQCEFLCLRLFCFSATRRYLGAVELSDVYQSDGELP

VDARPPVPSPPNPMNVYKKTLPVETHTLNIAAPRPMMGLLQPTQAEVYPT

VYITLKNASHPVSLHAYGVSYKASGAEYDDOTSOEKEDDDVPEGSHYVWVL

KENGMAADPLCLITYSLYSHVDLVKDNLSGLIALVCRSLAKKKTQTLAKLILF

AVFDEGSMHSEPTKNSLMQDDDAASARAPMKHNVGVNLSLGLGCHRSYWHY

IGMTPEVHSLFEGHTFLVNRHROALSLSPITFLAQTALDGOFLFCHISH

OHDEAEVYKVDSCPEEPOLPMKNNEEAEDDDDLTSEMVMVPPDDNSPFIOTS

VAKKPTWVYIAAEEDMDYAPLYLAADRSYKSOVLNRPORIGKTKYKVFMAV

TDEFTKREATHOHEGLIPLLIGEVGPTLLITIKNSRYNLTPIHDTIDVRLISL

RLPKGVKLKDEPILPGEIFKRYTVVDEGPTSDPEQLRTYRYSSENMEDRLASL

IGPLLCYKESVDQNGIMSDKRVNLLFVSDNRSVLYLENTLORELPPNAGVOLD

PEFASNIMSHINGVDFDLSOLVHEAVYLLISGAQDFLSEFSGYTFKHRY

YEDTLTLPFSGEYVFMENPGLMLCHNSDPRNMTALIVSSCDKTKGYED

SYEDTALLSKNNAIEPRSPNSRHSTKQFNATTPENDIEKTDPEARTPT

PKIONVSSDLMLROSPTPHGLSLDQAKKETSDDSPRAIDSNNSLSMTFR

RQQLHSGDMVTPESGLQTLNKLKTTATETLKLDFKVSSTSNMLISTIPDNL

AGTDNTSLGPPSPVHYDQDLDTLPEKSSPLTEGSGPLSEENNDKSLSGLM

NSOESWGRNVSSTEGRLFKGRHAGPALTKDNALFKVLSILKTNKSNSTNR

KTHIDGPTLLIENSFWONLLESTDEKRYTPIHMLMDKNATLRLHNNKKT

SSKNMEVQOKEGPIPPDAQNPMSPFKMLFLPESARWIDORTGKNSNGOCPSK

OLVSLGPEKSVGEONPLSEKRVVVGCEGPEKDKEMRPPSRNLPTINDLHN

NTNHEKTOEIEKEKTLIDENVLPDIHTVYTKNPMKMLFLSTRONEGSDGA

YAPVLDQRSRLNDSTNRTKHTAHFSKGEENLELGNDKQIVERYACTRTISPT

SOQNVYDQRSKRLKQRLPLETELEKRIYDQTSQMSKNMHLTPSLTQIDYNE

KEGATIQSPSLDCLTRSHSPIQANRSLPIAKTSFSPISPIYLTFLVLDONSHP

AASYRKDSGVSESHFLOGAKNNLSLALITLMTGDQREVGLGSATNSVYKYV

ENTVLPKPDLPKTSKGYELLPKVHLYOKDLPPTNSGSPCHDLVGSILQEGAI

KYNEANRPGKVPFLRATESSAKTPSKLLDPLDANHTGTQIPKEWKSQKSEKTA

FKKDDTILSLACSNHAIIAINEGQNKPELEIWMAROGTERLCSQNPVLRKHRE

ITRTTLOSDOEIYDDTISEMKKEPDIDEDENQSPRSFKTRHYFLAVERLM

DYGMSSPHVLRNRAQSGVPOPKVYVQETDGSFPOLYRGELNHLGLGYIRA

EVEDINWTFENRQASRPYSFSLISYEDROGCAFRKNVKNENKTYFWYOHNM

APTDEPCKAMAYESDYLKDYVSHGLIGLILCHNTLNPAAGROYVQERLFT

IFDEKSWYFTENMERNCRAPNTOMEDPTFKENRHHANGVIMDVLPGVMAQDOR

IRWYLLSGSNENHSHIFSGHVTYKARKKEEYKALNLYPGEVETVMPKSLAIR

VECLIGELHAGMSTFLVYSNKCOTPLGASGHIHIFOLATASQYOMAPKSLAIR

SGSINASTKEPESMIVKDLAPMIIGITQGRKOFKSLISYIOFLTNLSLOKMO

TYRNGSTGLTFVFGVNDGSGIKHNINPPIALYILPHLPHYSIRTLTMEIACDQ

NCSMPLGMEKASIDAOITASSFTYMTFRATWSKARHLQGRSNAMROYNPKPM

LQVDFQTKMKTGYTTOGVKSLTSMYKKEFLISSVDGQHWMTLFPNGKRVKVFQNG

Query Match 59.1%; Score 590.6; DB 9; Length 8967;
 Best Local Similarity 96.2%; Pred. No. 2, 5e-154;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy	372	actgacatcacctttctttctccacagatcatgcatccacatgcaaatagagctc	431
Db	66	ATTAACTTTGCTTCCACAGTGTGACATATTGTACATATAGCATCAATAGAGCTC	125
Qy	432	ccactgtctttctgtgaccttttgatcttgcttagtggccacgaagatactacc	491
Db	126	CCACTGCTTCTTCTGTGCTTGTGCAATTTCTCTTAGGCCACAGAAATATCTAC	185
Qy	492	tgggtcagtgtaactgtcagtgactatgcaaaagtatcctcgtgagctcgtg	551
Db	186	TGGGTGCGAGTGTGACTGTGATGAGGACTATATGCAAAATGATCTGTGACTGCTG	245
Qy	552	agcgaagatctctcctagagtgcaaaatctttccatccaacacctcagtcgtaca	611
Db	246	ACGCAAGATTTCTCTAGAGTGCCAAATCTTTCCATTCACACCTCAGTGTGTACA	305
Qy	612	aaaagactctgtgtgaattcagatcagatcactttcaaatcgtcgaagcagccac	671
Db	306	AAAAGACTCTGTGTGATGAAATTCACGAGATCACCTTTCAACATCGTAAAGCCAC	365
Qy	672	cctgagatggctcgtcaggtcctacatccaagcgtgaggttaataagatgctaca	731
Db	366	CCTGATGGGTCTGCTAGTGTCTACATCCAGGCTGAGGTTTATGATACAGTGTCTA	425
Qy	732	caactaagaacatgagctccatccatcgtcagctctcagctcgtgtgtgtactactga	791
Db	426	CACCTAAGAACATGCTGCTCCATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTG	485
Qy	792	aagcttcgaagagcgtgaatataatgaatgaatgaatgaatgaatgaatgaatga	851
Db	486	AAGCTTCGAGGAGCTGATATGATGATGATGATGATGATGATGATGATGATGATG	545
Qy	852	aagcttcctcgtggaagcgaatataatgctcgtcagctcgtcgaagagatgctcaa	911
Db	546	AAGCTTCCTCGTGGTGAAGCATATGATGATGATGATGATGATGATGATGATGATG	605
Qy	912	tggcctcgtgacacacgtgctcctacacatcattcttctcatgtgagctgtaaa	971
Db	606	TGGCCTCTGACCCCTGCTGCTTACCTACCTATATCTTCTCATGTGACCTGTAAAG	665
Qy	972	acttgaattcagagcctcattggaagccta 1000	
Db	666	ACTTGAATTACGAGCTCATTTGAGGCTTA 694	
RESULT	12		
LOCUS	HUMFVIII		
DEFINITION	Human coagulation factor VIII:C mRNA, complete cds.		
ACCESSION	M4113		
VERSION	M4113.1 GI:182817		
KEYWORDS	coagulation factor VIII:C.		
SOURCE	Human kidney, cDNA to mRNA, clones p8[100,102]; and cell line GM1416 DNA, clone p8-4.3.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 9029)		
AUTHORS	Truett, M.A., Blacher, R.W., Burke, R.L., Caput, D., Chu, C., Dina, D., Hartog, K., Kuo, C.H., Mastarz, F.R., Merryweather, J.P., Najjarian, R., Pachl, C., Potter, S.J., Puma, J., Quiroga, M., Rall, L.B., Randolph, A., Urdea, M.S., Valenzuela, P., Dahl, H.-H.M., Favalaro, J., Hansen, J., Nordfang, O. and Ezban, M.		
TITLE	Characterization of the polypeptide composition and expression of the human VIII:C and the nucleotide sequence and expression of the human kidney cDNA		
JOURNAL	JOURNAL		
MEDLINE	86081164		
REFERENCE	2 (sites)		

BASE COUNT 2841 a 1898 c 1833 g 2395 t
 ORIGIN

polyA-site

/note="polyA site"

misc_feature

/note="3' untranslated region"

misc_feature

/note="polyadenylation signal"

mat_peptide

/product="mature factor VIII (aa 1-2332)"

SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	1 (bases 1 to 9009)
AUTHORS	Lollar,J.S.
TITLE	Modified factor VIII
JOURNAL	Patent: US 6180371-A 1 30-JAN-2001;
FEATURES	Location/Qualifiers 1..9009
source	/organism="unknown"
BASE COUNT	2853 a 1907 c 1844 g 2405 t
ORIGIN	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:33:41 ; Search time 343.96 Seconds
(without alignments)
2492.514 Million cell updates/sec

Title: US-09-740-211-14_COPY_1_1000
Perfect score: 1000
Sequence: 1 cgcggccctcagcagcagctgcg.....cagcgcctcatgagagcccta 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_1101.*
1: /SID52/gcgdata/geneseq/geneseq/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	4899	21	Recombinant adeno
2	1000	100.0	4899	22	Human factor VIII
3	736.2	73.6	11933	21	Recombinant adeno
4	736.2	73.6	11933	22	Human factor VIII
5	640	64.0	7944	22	Plasmid DLZ6 encod
6	593.2	59.3	12445	21	Factor HSQ9Neo fo
7	592.2	59.2	5035	18	Factor VIII-dB695-
8	590.6	59.1	4832	19	Human factor VIII
9	590.6	59.1	4832	19	Human factor VIII
10	590.6	59.1	8967	17	Factor-VIII full-1
11	590.6	59.1	8967	22	Human factor VIII

12	590.6	59.1	8975	6	AAAS0054	Human factor VIII
13	590.6	59.1	8975	22	AAZ38604	Human full-length
14	590.6	59.1	9029	21	AAAF60309	Human factor VIII
15	590.6	59.1	9068	19	AAV15359	Human factor VIII
16	590.6	59.1	9080	19	AAV15580	Human factor III e
17	589	58.9	9009	14	AAAS0185	Human factor VIII
18	589	58.9	9009	18	AAAF61548	Factor VIII:C (Arg
19	589	58.9	9009	19	AAV25810	Human factor VIII
20	589	58.9	9009	20	AAAX91162	Human factor VIII
21	589	58.9	9009	22	AAAC90508	Human factor VIII
22	587.2	58.7	9164	20	AAAX82259	Beta-domain delete
23	587.2	58.7	11846	20	AAAX82261	Factor VIII protei
24	587.2	58.7	12022	20	AAAX82260	Factor VIII protei
25	586	58.6	6300	17	AAAT03571	Factor VIII CDNA.
26	586	58.6	8241	9	AAAB1439	Factor VIII CDNA.
27	586	58.6	8241	9	AAAB1096	CDNA sequence enco
28	585.8	58.6	9009	19	AAV18884	Homo sapiens facto
29	585	58.5	7053	18	AAAT51357	Factor VIII:C codi
30	584.2	58.4	7440	9	AAAB0689	Sequence encoding
31	583.4	58.3	4830	9	AAAB1544	Human factor VIII-
32	583.4	58.3	5094	21	AAAA49231	DNA construct HSQ/
33	583.4	58.3	7056	15	AAAO6615	Sequence of human
34	581.8	58.2	4629	20	AAAX88293	Human factor VIII
35	580.2	58.0	4629	16	AAQ76016	B-domain deleted F
36	579.2	57.9	9354	18	AAAT73164	CDNA encoding huma
37	566.6	56.7	7440	6	AAAB0375	DNA sequence enco
38	565.8	56.6	4373	20	AAAX82258	Beta-domain delete
39	528	52.8	4275	9	AAAB0447	Modified factor VI
40	528	52.8	4275	9	AAAB0446	Modified factor VI
41	528	52.8	4275	10	AAAB0654	DNA encoding factor VI
42	528	52.8	4545	9	AAAB0444	Human factor VIII
43	524.8	52.5	4670	19	AAV23339	Homo sapiens facto
44	521.8	52.2	9009	19	AAV12112	Plasmid DLZ7 encod
45	495.4	49.5	7914	22	AAAF84648	

ALIGNMENTS

RESULT 1	
ID AAD00122 standard; DNA: 4999 BP.	
XX AAD00122;	
AC	
XX	
DE 31-JUL-2000 (first entry)	
XX	
DE	Recombinant adeno associated vector construct, pYm4.lcF8deltaB.
KW	Recombinant Adeno Associated Vector; rAAV; pYm4.lcF8deltaB; promoter;
KW	human Factor VIII; hFVIII; EFlaIpha; human elongation factor-IaIpha;
KW	human growth hormone; hGH; IRR; inverted terminal repeat; haemophilia;
KW	gene therapy; ds.
XX	
OS	Adeno associated virus.
XX	
PN	W0200023116-A1.
XX	
PD	27-APR-2000.
XX	
PF	19-OCT-1999; 99WO-US24495.
XX	
PR	20-OCT-1998; 98US-0104994.
PR	24-MAR-1999; 99US-0125974.
XX	
PA	30-JUL-1999; 99US-0364862.
XX	
PI	(AVIG-) AVIGEN INC.
XX	
PI	Couto LB, Colosi PC;
XX	
DR	WPI: 2000-339536/29.
XX	
PT	New recombinant adenovirus-associated vector, useful for gene therapy

OY	195	tacttgctccgcgcctcttttcccggaaggttggtgggaagcaaccgtatataaagtgcagtgtgcc	250
Db	195	caacggccctgtgcctgtgccaaagtgtttctgtcgaacgaaccccaactgtgcctgtggcctgtgc	254
OY	255	gtcgaacgtctctttctgcgaacaggggttttcgcgcgccgcgcgcgcgaagtgcgaaggaatgt	314
Db	255	ataggccatcagcgaatgcgatctcagtggtgtgttttgcaagaggaagcaaaaagcctct	314
OY	315	tgtgtcttaataaacatctactga	370
Db	315	ccaccagagcctgtgaatgtttccacccaatgtcgaacagtggtgttttgcaagagaagc	374
OY	371	caactgaatccacgttttctctttcttccacaggtgatcatccaccatgcgaatatagatgc	430
Db	375	aaaaagcctctccaccacagcgcctgtgactcgaagatcttcgaaccacatgtcaataatagatctc	434
OY	431	tccaccctgtctcttctgtgtgcttttgcgattctgcgttttagtgcaccacgaatactac	490
Db	435	tccaccctgtctcttctgtgtgcttttgatctgtctttagtgcaccacgaatactac	494
OY	491	ctgggttgcaatgtggaactgtcaatggactatatagtcaaaagtatctcgtgtgagctgtgc	550
Db	495	ctgggttgcaatgtggaactgtcaatggactatatagtcaaaagtatctcgtgtgagctgtgc	554
OY	551	gacgcgaagaattctccctcagaatgtccaaatctttccattcaacacctgaatcgtgtac	610
Db	555	gacgcgaagaattctccctcagaatgtccaaatctttccattcaacacctgaatcgtgtac	614
OY	611	aaaaagactcgtttgtagaatccaagatcaaccttttcaacatcgtctaaagccaagcca	670
Db	615	aaaaagactcgtttgtagaatccaagatcaaccttttcaacatcgtctaaagccaagcca	674
OY	671	ccctgtgagtggtctgtccagagtcctcaccatcagctgtgaagttatgatcaagtgatcatt	730
Db	675	ccctgtgagtggtctgtccagagtcctcaccatcagctgtgaagttatgatcaagtgatcatt	734
OY	731	aacacttaagaacaatgtgtctccatcctgtcagttcctcatatgtctgttggtatcctactgg	790
Db	735	aacacttaagaacaatgtgtctccatcctgtcagttcctcatatgtctgttggtatcctactgg	794
OY	791	aaagctcttgaggagagctgaatatgatcatcagcaacagttcaaaaggagaaagaagaatgat	850
Db	795	aaagctcttgaggagagctgaatatgatcatcagcaacagttcaaaaggagaaagaagaatgat	854
OY	851	aaagctctccctgttggaagccaataatgtctgcgaagtccttgaaagaagaatgtcca	910
Db	855	aaagctctccctgttggaagccaataatgtctgcgaagtccttgaaagaagaatgtgtcca	914
OY	911	atggcctctgaaccaactgtgtccttaactatcatcttctcatatgttgaccctgtgtaaa	970
Db	915	atggcctctgaaccaactgtgtccttaactatcatcttctcatatgttgaccctgtgtaaa	974
OY	971	gacttgaattcagagcctcatgtgaagcccta	1000
Db	975	gacttgaattcagagcctcatgtgaagcccta	1004

RESULT	6
AAAA9232	
ID	AAAA9232 standard; DNA: 12445 BP.
XX	
AC	AAAA9232;
XX	
DT	26-SEP-2000 (first entry)
XX	
DE	Vector HSQreNeo for transforming endothelial cells.
XX	
KW	Hemostatic; antiplatelet; antidiabetic; neurotrophic; neuroprotective; vector
KW	osteoplastic; anticracking; immunostimulant; gene therapy; collagen;
KW	endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;
KW	vascular endothelial growth factor; bovine brain extract; haemophilia;
KW	Factor VIII; human; transgene; adenosine deaminase deficiency; ss;

KW	Sickle cell anaemia; thalassemia; diabetes; alpha-antitrypsin deficiency;
KV	Alzheimer's disease; brain disease; heart disease; immune system defect;
KW	bone fracture; osteoporosis.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
PN	WO200032750-A1.
XX	
PD	08-JUN-2000.
XX	
PF	24-NOV-1999; 99WO-US28033.
XX	
PR	24-NOV-1998; 98US-0109687.
XX	
PA	(MNU) UNIV MINNESOTA.
PA	(UYEM-) UNIV EMORY.
PA	(HEBH/) HEBBEL R P.
PA	(LINY/) LIN Y.
PA	(LOLL/) LOLLAR J S.
XX	
PI	Hebbel RP, Lin Y, Lollar JS;
XX	
DR	WPI: 2000-412303/35.
XX	
PT	Expanding population of endothelial cells useful to biocompatibilize
PT	implantable medical devices comprises contacting buffy coat cells with
PT	collagen I coated surface in culture medium comprising vascular
PT	endothelial growth factor -
XX	
XS	Claim 19, Fig 4; 53pp; English.

CC The invention relates to a method for expanding the population of
CC endothelial cells (EC) obtained from peripheral blood, by culturing,
CC in contact with a collagen I coated surface, buffy coat cells obtained
CC from peripheral mammalian blood in the presence of a culture medium
CC containing vascular endothelial growth factor (VEGF) and free of bovine
CC brain extract. EC are useful for treating hemophilia by introducing it
CC into the blood stream of a mammal, so that an effective amount of
CC Factor VIII protein is secreted in the blood stream of the mammal.
CC This sequence represents the vector HSOReko where the human factor
CC VIII gene into which an enhanced green fluorescent protein coding
CC sequence (HSG) has been inserted, is subcloned. Transgenic EC transduced
CC in vitro are useful for improving prosthetic implants. EC is also useful
CC for diagnosing clotting disorders where indication or disease is
CC associated with a reduction in the activity of an enzyme. EC is also
CC useful in gene therapy for treating the variety of diseases including
CC adenosine deaminase deficiency, sickle cell anemia, thalassemia,
CC hemophilia, diabetes, alpha-nitrylpsin deficiency, brain disorders
CC such as Alzheimer's disease, heart diseases, defects in immune system,
CC for repairing bone fractures and to treat or prevent osteoporosis.

Query Match	59.3%;	Score 593.2;	DB 21;	Length 12445;
Best Local Similarity	97.1%;	Pred. No. 8.6e-162;		
Matches 604;	Conservative 0;	Mismatches 18;	Indels 0;	Gaps 0

QY	379	tcacacttctctttcttccacagatctgataccacatgcgaatctagagctctccactg	438
Db	731	tcgcgatgcaccggatcgcgaaacctctctgcagccaccatgcgaatatagctctccactg	790
QY	439	ctctcttctctgaccttctgcgatctgcgtcttctagctgcaccagaatactacactctgctgc	488
Db	791	ctctcttctctgaccttctgcgatctctctcttctagctgcaccagaatactacactctgctgc	850
QY	499	agtgcgaactcatatgcagactatgcacaaatgtatcttcggttgagcgcctctgtagcgaag	558
Db	851	agtgcgaactctgcatagcgacctatgcacaaatgtatcttcggttgagcgcctctgtagcgaag	910
QY	559	attctctctcatgagtgcacaaatcttctccatccaaacacttagctcgtgtatcaaaaagac	618

Db	911	atttcctccctcaggctgicccaaatatcttccatccaacccctcagctcgtgatacaaaaagac	970
QY	619	tctgttcttgtaaatcaatcaagatcaacctttccaactgcgtcaagccaaagccacctggtat	678
Db	971	tctgttttgtaagaattcaagcttcaaccttttccaacatcgtcctaagccaaagccacctggtat	1030
QY	679	gggtcgtcgtcaggtgcctcaacatccagcctgaggtttatatacaagtggtcattcaacttaa	738
Db	1031	gggtcgtcgtcaggtgcctcaacatccagcctgaggtttatatacaagtggtcattcaacttaa	1090
QY	739	gaacatgctctcccatcctcgttcagcttcttcaatgctgttggtatcctcactcgtgaagcttc	798
Db	1091	gaacatgctctcccatcctcgttcagcttcttcaatgctgttggtatcctcactcgtgaagcttc	1150
QY	799	tgaaggaagctcgaatatgatatgaatcagaagccagctcaaaaggaggaagaagtgataaagtctt	858
Db	1151	tgaaggaagctcgaatatgatatgaatcagaagccagctcaaaaggaggaagaagtgataaagtctt	1210
QY	859	ccctcgtcgtgaaagccatacatatgctcgtgaaagttcctcgtgaagaagaaatgttccaatgcttc	918
Db	1211	ccctcgtcgtgaaagccatacatatgctcgtgaaagttcctcgtgaagaagaaatgttccaatgcttc	1270
QY	919	tgaaccacgtgtgcttaacctactacatatacttctcattgttgscctgtgtaaaagacttgaa	978
Db	1271	tgaaccacgtgtgcttaacctactacatatacttctcattgttgscctgtgtaaaagacttgaa	1330
QY	979	ttaagcgtcattatggagcccta 1000	
Db	1331	ttaagcgtcattatggagcccta 1332	

RESULT	7
AAT69811	
ID	AAT69811 standard; DNA; 5035 BP.

AC AAT69811;

DT 10-AUG-1997 (first entry)

Factor VIII-dB695-HCII DNA.

KW Factor VIII-db695-HCII; heparin cofactor II; blood coagulation;

KW haemophilia; gene therapy; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
25	5030	

```
FT      / *tag = a
FT      /mejust = Doctor Witt-dm605-witt
```

```
ET misc_RNA 2225..2314
DE /+++= b
```

/product= heparin cofactor II region (aa51-81)

PN W09718315-A1.

PD 22-MAY-1997.

PF 13-NOV-1996; 96WO-EP04977.

PR 13-NOV-1995; 9505-0558107.

PA (IMMO) IMMUNO AG.
YY

PI Voorberg JJ;
YY

DR WPI; 199/-289291/26.
DR P-PSDB: AAW18670

XX	Hybrid Factor VII
DT	

PT donor anticoagulant or antithrombotic protein - useful for treatment of coagulation disorders

XX Claim 16: Page 52-60; 96bp; English.

PS

XX A DNA molecule (AAT69811) codes for Factor VIII-dB695-HCII (AAW18670),

CC a hybrid protein in which amino acids 712-736 of Factor-dB695

CC (Factor VIII del868-1562) B-domain are replaced by amino acids

CC 51-80 from the acidic region (and potential thrombin-binding site)

CC of human heparin cofactor II (HCII). It was obt'd. by PCR

CC amplification (see also AAT69812-13) of the HCII acidic region from

CC total liver cDNA, fusion to sequences encoding Factor VIII

CC aa706-711 and aa737-743, and incorporation of the construct into

CC plasmid pCLR-dB695. The hybrid protein, which can be expressed

CC using gene therapy techniques, has increased procoagulant activity

CC owing to the HCII acidic region, and can be used to treat blood

CC coagulation disorders such as haemophilia A.

XX

XX Sequence 5035 BP; 1484 A; 1127 C; 1110 G; 1314 T; 0 other;

50

Query Match	59.2%;	Score 592.2;	DB 18;	Length 5035;
Best Local Similarity	97.9%;	Pred. No. 1.1e-161;		
Matches 600;	Conservative	0;	Mismatches 13;	Indels 0;
				Gaps 0;

QY	368	tctttttccacagatgtatgcataccatgatgaataagatccctccacctgtcttct	447
Db	7	tcagattgaacatttgttgcgaagccaccatggaataatagaagctctccactgtcttct	66
QY	448	gtgccttttgcgattctgcctttaigtgcacaagaatactactcctgtgtgcagttgaact	507
Db	67	gtgccttttgcattctcttattagtgccaccagaatactactcctgtgtgcagttgaact	126
QY	508	gtcacgggactatagcaagaatgactcgttgatgcctctgtgagccgcaagaatttccctc	567
Db	127	gtcacgggactatagcaagaatgactcgttgatgcctctgtgagccgcaagaatttccctc	186
QY	568	tagagtgcacaaatcttttccattcaaaccttcagtcgtgtacaaaagaactcttttgt	627
Db	187	tagagtgcacaaatcttttccattcaaaccttcagtcgtgtacaaaagaactcttttgt	246
QY	628	agaattcagagatcacaccttttloacatcgtctlaagccaaaggccacctgtatgtgtcgt	687
Db	247	agaattcagagatcacaccttttloacatcgtctlaagccaaaggccacctgtatgtgtcgt	306
QY	688	agtgcttaacatccagcggtgcaggttattatatacaagtgatcattacctaagaacatgac	747
Db	307	agtgcttaacatccagcggtgcaggttattatatacaagtgatcattacctaagaacatgac	366
QY	748	ttcccatcctgtcagttcttcatcgtctgtgtgvtalctactctggaagctcttgaaggagc	807
Db	367	ttcccatcctgtcagttcttcatcgtctgtgtgvtalctactcttgaagctcttgaaggagc	426
QY	808	tgaatatgtatgtcagcaagcagtcacaaaggagaaagaagtgtataaagctcttccctgttg	867
Db	427	tgaatatgtatgtcagcaagcagtcacaaaggagaaagaagtgtataaagctcttccctgttg	486
QY	868	aagccatacatatgtctgtgcaggtccctgcaagaagaatgttccaaatgtgcctctgaaccact	927
Db	487	aagccatacatatgtctgtgcaggtccctgcaagaagaatgttccaaatgtgcctctgaaccact	546
QY	928	gtgccttacctactaatatcttctctcatgttgaaacctgtaaaagaacttgaattcagagct	987
Db	547	gtgccttacctactaatatcttctctcatgttgaaacctgtaaaagaacttgaattcagagct	606
QY	988	caattgagcccta 1000	
Db	607	catltagagcccta 619	

RESULT	8
AAV19581	
ID	AAV19581 standard; CDNA; 4832 BP
XX	
AC	AAV19581;

(CHIR) CHIRON CORP.
 Allen JR, Barber JR, Boder M, Chang SM, Chong K;
 De La Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;
 Jolly DJ, Mittelstaedt DM, Prussak CE, Respass JD;
 WPI; 1998-086967/08.
 P-PSDB; AAM4372.
 New replication defective recombinant retroviruses - which express B
 domain-deleted human factor VIII or human factor IX for the
 treatment of haemophilia
 Claim 6; Page 174-175; 236pp; English.
 This DNA sequence includes a coding region for the B domain
 deletion mutant SON (see AAM4372) of human Factor VIII. The SON
 mutant is created by fusing Ser-743 to Gln-1638 of native Factor
 VIII (see AAM4373) to form a Ser-Gln-Asn (SON) link between the A2
 and A3 factor VIII domains. When compared to plasmid-derived
 Factor VIII, the SON deletion does not influence the in vivo
 pharmacokinetics, but the reduced size of the molecule appears to
 decrease proteolytic degradation. The invention relates to
 preparations of replication defective recombinant retrovirus (RV)
 expressing a B domain-deleted human Factor VIII protein, where the
 recombinant RV is capable of infecting human cells, is resistant
 to degradation by human complement and is capable of inducing
 long-term (at least 30 days and up to 6 months or longer
 post-injection) systemic expression of Factor VIII when
 administered to a haemophilia A patient.
 Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;
 Query Match 59.1%; Score 590.6; DB 19; Length 4832;
 Best Local Similarity 96.2%; Pred. No. 3.2e-161;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 912 tggcctctgaccacacgtctgcttactactatcatcttctcatgtgaccctgtaaaag 971
 DB 568 tggcctctgaccacacgtctgcttactactatcatcttctcatgtgaccctgtaaaag 627
 QY 972 actgaatcagcctcatgtgagacctta 1000
 DB 628 actgaatcagcctcatgtgagacctta 656
 RESULT 10
 AAT31031
 ID AAT31031 standard; cDNA; 8967 BP.
 AC AAT31031;
 DT 19-SEP-1996 (first entry)
 DE Factor-VIII full-length cDNA.
 KW Factor-VIII; retrovirus; vector; haemophilia A; gene therapy; ds.
 OS Homo sapiens.
 FX
 FT Key Location/Qualifiers
 FT CDS 110..7165
 FT /tag= a
 FT sig_peptide 110..166
 FT /tag= b
 FT mat_peptide 167..7162
 FT /*tag= c
 PN WO9621035-A2.
 PD 11-JUL-1996.
 PF 18-DEC-1995; 95WO-US16582.
 PR 30-DEC-1994; 94US-0366851.
 PA (CHIR) CHIRON VIAGENE INC.
 PI Bodner M, Chang S, Chi-Tang Hsu D, De Polo NJ;
 DR WPI; 1996-334010/33.
 DR P-PSDB; AAM00465.
 PT Retroviral vector directing expression of full length factor VIII -
 used in the gene therapy and treatment of haemophilia A
 PS Claim 3; Page 58-68; 86pp; English.
 XS A full-length cDNA clone (AAT31031) codes for human Factor-VIII
 CC (AAM00465), a trace plasma glycoprotein which acts as a cofactor in
 CC conjunction with Factor-IXa in the activation of Factor-X.
 CC Retroviral vectors comprising the full-length cDNA can be efficiently
 CC packaged into infectious retroviral particles. These may be used to
 CC transduce cells either in vivo or ex vivo. Factor-VIII expressed
 CC from such transduced cells will be processed and transported in a
 CC fashion analogous to the expression product of a normal Factor-VIII
 CC gene. Retroviral particles harbouring such vectors will be useful
 CC in the gene therapy of haemophilia A.
 QY Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;
 DB 372 actgacatcacatttcttcttctccacagatcatgacacacgaataagagctct 431
 DB 66 attacatttgcctccagttgaacatttgaacataagtcacgaataagagctct 125
 Query Match 59.1%; Score 590.6; DB 17; Length 8967;
 Best Local Similarity 96.2%; Pred. No. 4.2e-161;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY	432	caaccgcgtcttcttgctgctcttttggcgtatcttgctttagtgccaccgaagatactacc	491
Db	126	ccccccgctctctctctgctgctcttttggcgtatcttgctttagtgccaccgaagatactacc	185
QY	492	tgggttcagtggaactgcatactatgagacataatgacaagtatctcggtagctgcgtctgtg	551
Db	186	tgggttcagtggaactgcatactatgagacataatgacaagtatctcggtagctgcgtctgtg	245
QY	552	acgcgaagattctctcctagatggtccaaaactcttccatccaacactcgaatcgtgtaca	611
Db	246	acgcgaagattctctcctagatggtccaaaactcttccatccaacactcgaatcgtgtaca	305
QY	612	aaaagaactgtgttgaagattcacggatcaccttttcaacaatcgctaaagcgccac	671
Db	306	aaaagaactgtgttgaagattcacggatcaccttttcaacaatcgctaaagcgccac	365
QY	672	cctgtgaatgggtctgctagaagctccataccatccagctcgaagtttatatatacagtgtcat	731
Db	366	cctgtgaatgggtctgctagaagctccataccatccagctcgaagtttatatatacagtgtcat	425
QY	732	caactaagaacttgcttcccaactcctctgcagcttccatgctgttggtgtatcctactgga	791
Db	426	caactaagaacttgcttcccaactcctctgcagcttccatgctgttggtgtatcctactgga	485
QY	792	aagctctcagaagactgataatgatatgatacgaacgaatcgaagggagaagaagaatgata	851
Db	486	aagctctcagaagactgataatgatatgatacgaacgaatcgaagggagaagaagaatgata	545
QY	852	aagctctccctgtgtgaagccatacatatgtctcgtcgaagtcctcgaagaagatgtgccaa	911
Db	546	aagctctccctgtgtgaagccatacatatgtctcgtcgaagtcctcgaagaagatgtgccaa	605
QY	912	tggcctcagaccactgtgacctactactatcttccatcatttccatcagtgtgaactgtgtaaag	971
Db	606	tggcctcagaccactgtgacctactactatcttccatcatttccatcagtgtgaactgtgtaaag	665
QY	972	acttgaatcagagcctcatgtgaagcccta 1000	
Db	666	acttgaatcagagcctcatgtgaagcccta 694	
RESULT 11			
ID	AAC87526		
XX	AAC87526 standard; DNA; 8667 BP.		
AC	AAC87526;		
XX			
DT	13-MAR-2001 (first entry)		
XX			
DE	Human factor VIII cDNA, SEQ ID NO:1.		
XX			
KM	Factor VIII: human; A2 domain; C2 domain; LRP-mediated plasma clearance		
KW	receptor-dependent clearance; receptor-independent clearance;		
XX	haemophilia; half-life, ss.		
XX			
OS	Homo sapiens.		
XX			
PN	MO20007174-A2.		
XX			
PD	30-NOV-2000.		
XX			
PF	24-MAY-2000; 2000WO-US14111.		
XX			
PR	24-MAY-1999; 99US-0135847.		
XX			
FA	(AMNA-) AMERICAN NAT RED CROSS.		
XX			
PI	Saenko EL, Strickland DK;		
XX			
DR	WPI: 2001-025163/03.		
XX			
DR	P-PSDB; AAB48843.		
XX			

PT Factor VIII mutants having increased half-life useful for treating
 PT hemophilia, comprise one or more amino acid substitutions in the A2
 PT and/or C2 domain of factor VIII -
 XX
 XX
 PS Claim 41; Page 86-101; 121pp; English.
 CC The invention relates to human factor VIII mutants comprising an amino
 CC acid substitution at one or more positions in the A2 domain and/or an
 CC amino acid substitution at one or more positions in the C2 domain.
 CC The invention also encompasses a factor VIII mutant which lacks a B
 CC domain (A2B484). The factor VIII mutants have an increased half-life
 CC in the bloodstream. The A2 domain mutants exhibit reduced LRP-dependent
 CC (receptor-dependent) clearance of factor VIII, while C2 domain mutants
 CC have reduced receptor-independent clearance. The invention also relates
 CC to a method of using RAP (receptor associated protein), a protein which
 CC inhibits LRP (low density lipoprotein related protein)-mediated ligand
 CC internalisation, to increase the half-life of factor VIII. The mutant
 CC factor VIII proteins, and nucleotides encoding them, are useful
 CC for treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and
 CC nucleic acids encoding them may also be used in the treatment of
 CC haemophilia, in combination with a mutant factor VIII protein or DNA of
 CC the invention. The invention provides means of increasing the half-life
 CC of factor VIII by reducing its clearance from plasma. The present
 CC sequence represents cDNA encoding human factor VIII.
 XX
 XX Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;

Query Match	59.1%;	Score 590.6;	DB 22;	Length 8967;
Best Local Similarity	96.2%;	Pred. No. 4.2e-161;		
Matches 605;	Conservative	0;	Mismatches 24;	Indels 0;
				Gaps 0;

QY	372	actggaatccacaatcttcttcttcttcccaacgagtaatgataccacgaatgaagagctt	431
Db	66	atlaacoccttggctctcccaagttgaacatttgtgccaataagcaatagaagctt	1255
QY	432	ccacctgtctcttcttgctgaccttthgacattctgcttaagtgcaccaagaatactacc	4911
Db	126	ccacctgtctcttcttgctgaccttthgacattctgcttaagtgcaccaagaatactacc	1858
QY	492	tgggtgcaggtgaaactgtcattgaggactatagcaaatgatctcgtgtgagctcctgtg	5515
Db	186	tgggtgcaggtgaaactgtcattgaggactatagcaaatgatctcgtgtgagctcctgtg	2455
QY	552	acgcagaattcttcctcctgagtgccaaaattttccattcaacacactcagtcgtgtaca	6111
Db	246	acgcagaattcttcctcctgagtgccaaaattttccattcaacacactcagtcgtgtaca	3055
QY	612	aaaagaactcgtttgttagaatltaacggaatcaacctttcaacatcgtcaagccaaagccac	6711
Db	306	aaaagaactcgtttgttagaatltaacggaatcaacctttcaacatcgtcaagccaaagccac	3655
QY	672	cctggaatgggtcgtctgaagtccctacacaccaggtctgaagtttatgatacagtggtcaata	7211
Db	366	cctggaatgggtcgtctgaagtccctacacaccaggtctgaagtttatgatacagtggtcaata	4255
QY	732	caacttaagaacaatgggtctcccatcctgtlcaatcttcaatctgctgtgtgtgtatcccaactgga	7911
Db	426	caacttaagaacaatgggtctcccatcctgtlcaatcttcaatctgctgtgtgtgtatcccaactgga	4855
QY	792	aagctctctgaaggagactgaataatgatatcatcagaccagttcaaaaggagagaagaatgatata	8511
Db	486	aagctctctgaaggagactgaataatgatatcatcagaccagttcaaaaggagagaagaatgatata	5455
QY	852	aagttcttccctgtgtggaagccatacatatgtctggcaggtctcctgaaagagaatgtgtccaa	9111
Db	546	aagttcttccctgtgtggaagccatacatatgtctggcaggtctcctgaaagagaatgtgtccaa	6055
QY	912	tggcctctgaccacacatctgaccttaaccatacttcttcttaagtgtgagacctgttaaaag	9711
Db	606	tggcctctgaccacacatctgaccttaaccatacttcttcttaagtgtgagacctgtgttaaaag	6655
QY	972	acttgaattcagagccatcatatggagacccta	1000


```
Db 502 cacttaagaacatggtcccatccctcgtcagtcgtctcatgctgtgtatcctactgga 561
QY 792 aagcttcgagggagctgaatatga tcaagaccagtcagaaggagaagaagatgata 851
    |||
Db 562 aagcttcgagggagctgaatatga tcaagaccagtcagaaggagaagaagatgata 621
QY 852 aagcttcgccctggtggaagccatacatgtctgcaagtcctgaaagagaatggtccaa 911
    |||
Db 622 aagcttcgccctggtggaagccatacatgtctgcaagtcctgaaagagaatggtccaa 681
QY 912 tggcctcgaaccactgagccttaacctactcatcttctcaatgagacctggtaaaag 971
    |||
Db 682 tggcctcgaaccactgagccttaacctactcatcttctcaatgagacctggtaaaag 741
QY 972 acttgaattcaggcctcatgtgagcccta 1000
    |||
Db 742 acttgaattcaggcctcatgtgagcccta 770
```

Search completed: January 17, 2002, 16:34:09
Job time: 12554 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:27:13 ; Search time 118.22 Seconds
(without alignments)

1915.733 Million cell updates/sec

Title: US-09-740-211-14_COPY_1_1000

Sequence: 1 cgcacctgcagcagctgcg.....cagcctcatltgagacctta 1000

Scoring table:	IDENTITY_NUC	Canoe	100	Canoe	100
----------------	--------------	-------	-----	-------	-----

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing:	Minimum Match	0%
Minimum Match	10%	

Listing first 45 summaries

```
Database : Issued_Patents_NA:*
```

```

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB	ID	Description
1	1000	100.0	4999	4	US-09-470-618-14	Sequence 14, Appl
2	1000	100.0	4999	4	US-09-364-862-14	Sequence 14, Appl
3	736.2	73.6	11933	4	US-09-470-618-13	Sequence 13, Appl
4	736.2	73.6	11933	4	US-09-364-862-13	Sequence 13, Appl
5	592.2	59.2	5035	2	US-08-882-083-1	Sequence 1, Appl
6	592.2	59.2	5035	2	US-08-558-107-1	Sequence 1, Appl
7	592.2	59.2	5035	3	US-09-243-539-1	Sequence 1, Appl
8	590.6	59.1	8967	1	US-08-366-851A-1	Sequence 1, Appl
9	589	58.9	9009	1	US-07-864-004B-3	Sequence 3, Appl
10	589	58.9	9009	1	US-08-251-937A-3	Sequence 3, Appl
11	589	58.9	9009	1	US-08-212-133A-1	Sequence 1, Appl
12	589	58.9	9009	1	US-08-474-503-1	Sequence 1, Appl
13	589	58.9	9009	2	US-08-670-707A-1	Sequence 1, Appl
14	589	58.9	9009	4	US-09-037-601-1	Sequence 1, Appl
15	589	58.9	9009	5	PCT-US93-03275-3	Sequence 3, Appl
16	589	58.9	9009	5	PCT-US94-13200-1	Sequence 1, Appl
17	586	58.6	8241	6	5171844-1	Patent No. 5171844
18	583.4	58.3	7056	1	US-08-121-202-1	Sequence 1, Appl
19	581.8	58.2	4629	2	US-08-484-891-7	Sequence 7, Appl
20	579.2	57.9	9354	1	US-08-683-839B-2	Sequence 2, Appl
21	528	52.8	4670	3	US-08-717-294-41	Sequence 41, Appl
22	454.2	45.4	7032	4	US-08-276-584A-1	Sequence 1, Appl
23	451.4	45.1	7493	4	US-09-324-867-1	Sequence 7, Appl
24	451.4	45.1	7493	1	US-08-212-133A-7	Sequence 7, Appl
25	451.4	45.1	7493	2	US-08-474-503-5	Sequence 5, Appl
26	451.4	45.1	7493	4	US-08-670-707A-5	Sequence 5, Appl
27	451.4	45.1	7493	4	US-09-037-601-5	Sequence 5, Appl

28	451.4	45.1	743.5	5	PCr-US94-13200-5	Sequence 58, April
29	442.8	42.4	743.4	2	US-08-670-707A-38	Sequence 38, April
30	442.8	42.4	443.4	4	US-09-037-601-38	Sequence 38, April
31	442.8	42.4	640.2	2	US-08-670-707A-36	Sequence 36, April
32	442.8	42.4	640.2	2	US-09-037-601-36	Sequence 36, April
33	303	30.3	445.1	3	US-08-117-2294-42	Sequence 42, April
34	151.4	15.1	445.1	3	US-09-670-618-14	Sequence 14, April
35	151.4	15.1	499.9	4	US-09-664-862-14	Sequence 14, April
36	151.4	15.1	1193.3	4	US-09-670-618-13	Sequence 13, April
37	151.4	15.1	1193.3	4	US-09-664-862-13	Sequence 13, April
38	138	13.8	1185	1	US-07-757-538B-13	Sequence 5, April
39	138	13.8	1189	1	US-07-757-538B-13	Sequence 4, April
40	138	13.8	1561	1	US-07-757-538B-1	Sequence 1, April
41	138	13.8	5225348-3	6	5225348-3	Sequence 1, April
42	138	13.8	4695	6	5225348-3	Sequence 1, April
43	134.6	13.5	8295	1	US-08-662-014-2	Sequence 2, April
44	134.6	13.5	8295	1	US-08-923-137-3	Sequence 3, April
45	134.6	13.5	8295	4	US-08-973-334-5	Sequence 5, April

ALIGNMENTS

```

RESULT      1
US-09-470-618-14
Sequence 14, Application US/09470618
Patent No. 6200560
GENERAL INFORMATION:
APPLICANT: Couto, Linda B.
APPLICANT: Colosi, Peter C.
TITLE OR INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
TITLE OR INVENTION: by Target Cells
FILE REFERENCE: AVigen-04082
CURRENT APPLICATION NUMBER: US/09/470,618
CURRENT FILING DATE: 1999-12-22
EARLIER APPLICATION NUMBER: 09/364,862
EARLIER FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: 60/125,974
EARLIER FILING DATE: 1999-03-24
EARLIER APPLICATION NUMBER: 60/104,994
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 4999
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-14

```

Query Match	100.0%;	Score 1000;	DB 4;	Length 4999;
Best Local Similarity	100.0%;	Pred. No. 1.3e-305;		
Matches 1000;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	cgccctcgagagcagctctgcggtctgcgtcgtctcaactttagagccgcgccgggcaaaacccggg	60
Db	1	cgccccctcgagagcagctctgcgtctgcgtcctcaactttagagggccgcgccgggcaaaacccggg	60
QY	61	cgctcgcgcgaaacctttgtgtctgcgccgggacctcaagtgagcgagcgagcgcgcaagagagagctg	120
Db	61	cgctcgcgcgaaacctttgtgtctgcgccgggacctcaagtgagcgagcgagcgcgcaagagagagctg	120
QY	61	cgctcgcgcgaaacctttgtgtctgcgccgggacctcaagtgagcgagcgagcgcgcaagagagagctg	120
Db	61	cgctcgcgcgaaacctttgtgtctgcgccgggacctcaagtgagcgagcgagcgcgcaagagagagctg	120
QY	121	gccaaactcattactacaggggtttccgtgcgagcgcaacgcgctgtgtgcgcggggtaaacctggg	180
Db	121	gccaaactcattactacaggggtttccgtgcgagcgcaacgcgctgtgtgcgcggggtaaacctggg	180
QY	181	aaagagatctgcgtgtactcgtgctccgcacctttttcccgaggggtggggggagaaacggtataaa	240
Db	181	aaagagatctgcgtgtactcgtgctccgcacctttttcccgaggggtggggggagaaacggtataaa	240
QY	241	gtgcagctagtcgcgtcgtagaactttctttttctgcgaaggggttttcgcgcgcccgccggcgaggttaa	300
Db	241	gtgcagctagtcgcgtcgtagaactttctttttctgcgaaggggttttcgcgcgcccgccggcgaggttaa	300

[illegible]

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-364-862-14

```

Query Match	100.0%;	Score 1000;	DB 4;	Length 4999;
Best Local Similarity	100.0%;	Pred. No. 1.3e-305;		
Matches 1000;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 35..5017
 US-09-243-539-1

Query Match 59.2%; Score 592.2; DB 3; Length 5035;
 Best Local Similarity 97.9%; Pred. No. 8e-177;
 Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 388 tcttttccacaggtatccacccatgcaaatagagctctccactcttctt 447
 Db 7 TCCAGTTGAACATTGTAGACAGCCACATGGAATAGAGCTCTCCACTGCTTCTTCT 66
 QY 448 gtgcctttgagcttgccttttagtgcacccaggaagatactccctgggtgcaatggaact 507
 Db 67 GTGCTTTTGCATTTCTGCTTTTAGTGCACCCAGAAAGATACCTGCGGTGCAAGTGAAGT 126
 QY 508 gtcaatggaactatagcaaatgagctgcgtgagctgcctgctggaagaaattctcc 567
 Db 127 GTCATGGGACTATATSCAAAGTATCTCGGTGAGCTGCTGAGCAAGATTTCCCTCC 186
 QY 568 tagatggcaaaatctttccatcaaacctcagctgctgtaacaaagactctgttct 627
 Db 187 TAGAGTGCAGAAATCTTTTCCATTCACACCTCAGTCTGTAACAAAAGACTCTGTTGT 246
 QY 628 agaatcaagatcaaccttttcaacatcgttaagccaagccaacctgagatggctgct 687
 Db 247 AGAATTCAGAGATACCTTTTCAACATCGTAAAGCCACCTGAGTGGTCTGCT 306
 QY 688 aggtctaccatccagctgaggttatgatacagtgatcaacttaagaaacatgagc 747
 Db 307 AGGTCTTACCAATCCAGGCTGAGGTTATGATACAGTGTCTACATTAAGAACATGGC 366
 QY 748 ttcccatcctgcagctcttcaatcagctgctggtgtatccactggaagactctctggagc 807
 Db 367 TTCCTACTCTGCTCACTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
 QY 808 tgaatagatgatacagacagatcaaaaggaagaaatataaaatgcttccctgtg 867
 Db 427 TGAATATGATGATCAGACAGCAGTCAAAAGGAGCAAAAGATATATAAGTCTTCCCTGCTG 486
 QY 868 aagcatcatatgctgtgcagctgcctggaagaaatgctcaatgagctctgagccact 927
 Db 487 AAGCATCATATGCTGTGAGGCTGCTGAAAGAGAAATGCTCAATGCTGCTGCTGCTGCTGCT 546
 QY 928 gtgccttaccatcatatcttctcatgtagactggttaaaagacttgatcaagct 987
 Db 547 GTGCTTACATACATATCTTTCTCATGTGAGCTGCTGTAAGAACTTGAATTCAGGCT 606
 QY 988 cattgagacccta 1000
 Db 607 CATTGAGCCCTTA 619

RESULT 8
 US-08-366-851A-1
 Sequence 1, Application US/08366851A
 Patent No. 5681746

GENERAL INFORMATION:
 APPLICANT: Bodner, Mordechai
 APPLICANT: De Folo, Nicolas J.
 APPLICANT: Hsu, David Chi-Tang
 APPLICANT: Chang, Steven
 TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Viagene, Inc.
 STREET: 11055 Roselle Street
 CITY: San Diego
 STATE: California

COUNTRY: U.S.A.
 ZIP: 92121
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/366,851A
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Chambers, Daniel M.
 REGISTRATION NUMBER: 34,561
 REFERENCE/DOCKET NUMBER: 930049.438
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 452-1288
 TELEFAX: (619) 452-2616
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8967 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 110..7165
 US-08-366-851A-1

Query Match 59.1%; Score 590.6; DB 1; Length 8967;
 Best Local Similarity 96.2%; Pred. No. 3.6e-176;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 372 actgacatccacttttcttccacaggtatcgatccacatgcaaatagagctct 431
 Db 66 ATTAACCTTTTGTCTCTCAGTGAACATTGTGACATTAAGTCATGCAATGAGACTCT 125
 QY 432 ccactgcttcttctgtgcttctgtgcttctgtgcttctgtgcttctgtgcttctgtgctt 491
 Db 126 CCACCTGCTTCTTCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 185
 QY 492 tgggtgagtggaactgcatggaactatgcaaatgagtgagtcggtgagctgctg 551
 Db 186 TGGGTGAGTGGAAGTGTGATGAGGAGTATGCAAAAGTATGCTGCTGCTGCTGCTGCTGCT 245
 QY 552 acgcaagattctctctagagtgcaaatcttccatcaaacctcagctggttaca 611
 Db 246 ACGCAAGATTTCTCTCTAGAGTGCAGAAATCTTTTCCATTCACACCTCAGTGTGTACA 305
 QY 612 aaagactcgttctgtagaaltcaagatcaacttttcaaacctcgttaagccaagccac 671
 Db 306 AAAAGACTCTTTTGTGATTCACGGATCACCTTTTCAACATCGCTAAGCCAAAGCCAC 365
 QY 672 cctgagtggtctgtagtgcctaccatccagagctgaggttitaagtaagtgatca 731
 Db 366 CCGTATGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425
 QY 732 cacttaagaacatggtctccatccctgctcagcttcaatgctggttgggtatcctc 791
 Db 426 CACTTAAGAATGAGTCCCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 485
 QY 792 aagcttggagggagcgtgaatgataagcaagcaagcaagcaagcaagcaagcaagcaag 851
 Db 486 AAGCTTGTGAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545
 QY 852 aagcttccctggtggaagccacacatgctgaggttctggaaggaaggaaggaaggaag 911
 Db 546 AAGCTTCCCTGCTGAGGAGCATATATGATGATGATGATGATGATGATGATGATGATGATGAT 605
 QY 912 tggccttgacccaactgtgcttaactactatcttctcaatgtagactggttaag 971

Db 606 TGGCCTGACCCAGTGGCTTACTACTATATCTTTCTCATGTGAGCTGTGTAAG 665
Qy 972 actgaattcagccatcattgagacccta 1000
Db 666 ACTTGAAATTCAGGCTCATTTGAGACCCCTA 694

RESULT 9
US-07-864-004B-3

Sequence 3, Application US/07864004B
Patent No. 5364771

GENERAL INFORMATION:

APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/864,004B
FILING DATE: 07 April 1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU106
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6553

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 9009 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:

ORGANISM: Homo sapien
TISSE TYPE: Liver

FEATURE:

NAME/KEY: misc.feature (Domain Structure)
LOCATION: 5001 . . . 7053
OTHER INFORMATION: /note="Equivalent to the A3-C1-C2
OTHER INFORMATION: domain"
FEATURE:
NAME/KEY: misc.feature (Domain Structure)
LOCATION: 1 . . . 2277
OTHER INFORMATION: /note="Equivalent to the A1-A2
OTHER INFORMATION: domain"
US-07-864-004B-3

Query Match 58.9%; Score 589; DB 1; Length 9009;
Best local Similarity 96.0%; Pred. No. 1,1e-175;

Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 372 actgaatcacttttcttttctccacaagatcaccacatgcacaatagagctct 431
Db 107 ATTAACTTTGGCTTCAGTTGACATTTGTAGCAATTAAGTCATGCAATAGAGCTCT 166

Qy 432 ccaactgcttcttctgtgcttcttgagatctgctttagtgcacccaagatactacc 491
Db 167 CCAGCTGCTTCTTCTTCTGCTGCTTTCGAGTCTGCTTTAGTGGCACCAAGATTAACC 226
Qy 492 tgggtcagtgaaactgcaatggaactatatagcaagtgatcctggtgagctgtgtg 551
Db 227 TGGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 286
Qy 552 agcgaagatctcctccttagagtgccaaatcttctcattcaaacctgaactggtgaa 611
Db 287 ACCGAAGATTTCTCTCTAGAGTGCACAAATCTTTTCATTCAACACCTGAGTGTGAGA 346
Qy 612 aaaaagactcgttctgtagaattcaagatcaccttccaaatcgtgaaagcgaagccac 671
Db 347 AAAAGACTCTGTTGTGAGATTTACGGTTTCACCTTTTCAACATCGTTAAGCCAGGCCAC 406
Qy 672 cctgagatggtctgctagtgctcctacacacagcctgaggttagtatacagtggtta 731
Db 407 CCTGAGATGGGTCTGTAGTGTCTACCATCCAGGCTGAGGTTTATGATAGAGTGTCTATTA 466
Qy 732 caactaagaacatggtctccatcctgctgctgcttctgctgctgctgctgctgctg 791
Db 467 CACTTAAGAACATGGCTTCCATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
Qy 792 aagcttctgagagagctgaaatgatgatgatacagcagcgaaggaaggaagatgata 851
Db 527 AAGCTTTGAGGAGAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 586
Qy 852 aagcttccctgctgctgagagcacaatatactgctgcaagctcgaaggaagatgctcaa 911
Db 587 AAGCTTCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 646
Qy 912 tggcttctgagagagctgctgcttaccatcctacatcttctcattgagagctgtgtaa 971
Db 647 TGGCTTCTGAGCCAGCTGCTGCTTACCTTACCTATATCTTCTCATGTGAGCTGTGTAAG 706
Qy 972 actgaattcagccatcattgagacccta 1000
Db 707 ACTTGAAATTCAGGCTCATTTGAGACCCCTA 735

RESULT 10

US-08-251-937A-3
Sequence 3, Application US/08251937A
Patent No. 5583209

GENERAL INFORMATION:

APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/251,937A
FILING DATE: 31-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EMU106DIV

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6367
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
    LENGTH: 9009 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: cDNA
    HYPOTHETICAL: NO
    ANTI-SENSE: NO
    FRAGMENT TYPE: N-terminal
    ORIGINAL SOURCE:
    ORGANISM: Homo sapien
    TISSUE TYPE: Liver
FEATURE:
    NAME/KEY: misc_feature (Domain Structure)
    LOCATION: 5001..7053
    OTHER INFORMATION: /note="Equivalent to the A3-C1-C2
    OTHER INFORMATION: domain"
US-08-251-937A-3

Query Match          58.9%; Score 589; DB 1; Length 9009;
Best Local Similarity 96.0%; Pred. No. 1,1e-175;
Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 372 actgacatccattcttcttctccacaggtatccaccatgcaaatagagcct 431
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 107 ATTACCTTTTCTCTCTCAGTTGAACATTTTACCAATTAAGTCATGCAATAGAGCTCT 166
QY 432 ccaactgcttcttctgcttctgacattctgcttctgacacacagaaatactacc 491
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 167 CCACCTGCTTCTTCTGTCGCTTTGCGATCTGCTTAAAGCCAGGAAGATCTACCC 226
QY 492 tgggtgacgtggaactgtcatgagactatagcaagtgatcctggtgagctgctgtgg 551
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 227 TGGGTGACGTGGAAGTGCATGAGACTATATGCAAGATGATCTCGTACGCTGCTGTGG 286
QY 552 acgcaagatctcctcctcagatgacgaacaccttctcattcaaacaccctgctgtgaca 611
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 287 ACGCAAGATTTCTCTCTAGAGTGCACAAATCTTTCCATTCAACACCTCACTCGTGTACA 346
QY 612 aaagaactcgttctgtagaatcacgatacccttcaacatcgtaagcgaagccac 671
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 347 AAAAAGACTCTGTTGTAGAAATTCACAGGTTTCACCTTTCAACATGCGCTAAGCCAGGCCAC 406
QY 672 cctggatgggtcgtctagatgctcaccatcagcaggtgaggttatagatacagtgcatla 731
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 407 CCTGATGGGTTGCTAGGTCTTACATTCACAGGCTGAGGTTTATGATACATGCTCATTA 466
QY 732 cacttaagaacatggctcccatcctgcaagcttctcattcagctgtgtgtatcctcactgga 791
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 467 CACTTAAGAACAATGGCTTCCATCTGTCAGTCTTCAATGCTGTGGTGTATACCTACATGCA 526
QY 792 aagcttcgagggagctgtaataatgatgatcatcagacagatcaaggagagaagaatgata 851
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 527 AAGCTTTCAGAGGAGCTGAATATATGATGCAGACACAGCAAGGAGCAAAAGAAATGATATA 586
QY 852 aagcttcctcggtggagagacataatgtctcagaggtccttaagaagagatgtgtccaa 911
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 587 AAGTCTTCCCTGGTGGAGGACATATATGCTGTGCAAGGCTCTAAGAGAGATGTCCAA 646
QY 912 tggcctctgacccacatgtgcttaacctactactatcttctcattgtgagactggtlaaag 971
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 647 TGGCCTTGAGCCCACTGCTTACCTACTACATATCTTTCTCAATGTGGAGCTGTGTAAG 706
```

```
QY 972 actgaattcagagcctcattggagcccta 1000
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 707 ACTGAATTTCAGGCTCTATTGGAGCCCTTA 735

RESULT 11
US-08-212-133A-1
Sequence 1, Application US/08212133A
Patent No. 5663060
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,133A
FILING DATE: March 11, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-Apr-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU/76677
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
    LENGTH: 9009 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: cDNA
    HYPOTHETICAL: NO
    ANTI-SENSE: NO
    ORIGINAL SOURCE:
    ORGANISM: Homo sapien
    TISSUE TYPE: Liver
FEATURE:
    NAME/KEY: misc_feature (Domain Structure)
    LOCATION: 5125..7053
    OTHER INFORMATION: /note="Equivalent to the A3-C1-C2
    OTHER INFORMATION: domain"
FEATURE:
    NAME/KEY: misc_feature (Domain Structure)
    LOCATION: 1..2277
    OTHER INFORMATION: /note="Equivalent to the A1-A2 domain.."
FEATURE:
    NAME/KEY: Domain
    LOCATION: 1..2277
    OTHER INFORMATION: /note="CDNA encoding human factor
    OTHER INFORMATION: VIII."
US-08-212-133A-1

Query Match          58.9%; Score 589; DB 1; Length 9009;
Best Local Similarity 96.0%; Pred. No. 1,1e-175;
Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
```

```

QY 372 actgacatcaccttcttcttcttccacagtgatcatccatcgaatagagctc 431
DB 107 ATTAACCTTTGCTTCCAGTGAACATTTGATGACATAGTATGCAATAGAGCTCT 166
QY 432 ccactgtcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 491
DB 167 CCACCTGCTTTCTGCTGCTTGGGATTCCTTTAGGCCACCAAGAAAGTACTACC 226
QY 492 tgggtgcagtgaaatcgtatcagtgaaatcgtatcagtgaaatcgtatcagtg 551
DB 227 TGGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 286
QY 552 acgcaagattcttcttcttcttcttcttcttcttcttcttcttcttcttctt 611
DB 287 ACGCAAGATTCTCTCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 346
QY 612 aaagacatcttcttcttcttcttcttcttcttcttcttcttcttcttctt 671
DB 347 AAAAGACTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 406
QY 672 cctggtgagtgctgctgctgctgctgctgctgctgctgctgctgctgctg 731
DB 407 CCGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 466
QY 732 cacttaagaacatgctgctgctgctgctgctgctgctgctgctgctgctgct 791
DB 467 CACTTAAGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 526
QY 792 aagcttcgagagagctgctgctgctgctgctgctgctgctgctgctgctgct 851
DB 527 AAGCTTCGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 586
QY 852 aagcttcctggtgagagagagagagagagagagagagagagagagagagag 911
DB 587 AAGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 646
QY 912 tggctctgagagagagagagagagagagagagagagagagagagagagag 971
DB 647 TGGCTCTGAGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 706
QY 972 actgaattcagagcctcatctgagagcctca 1000
DB 707 ACTGAATTCAAGGCTCATGATGAGAGCCCTA 735

RESULT 12
US-08-474-503-1
; Sequence 1, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106CIP(3)

```

```

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9009 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver
; FEATURE:
; NAME/KEY: misc-feature (Domain Structure)
; LOCATION: 5125..7053
; OTHER INFORMATION: /note="Equivalent to the A3-C1-C2
; OTHER INFORMATION: domain"
; FEATURE:
; NAME/KEY: misc-feature (Domain Structure)
; LOCATION: 1..2277
; OTHER INFORMATION: /note="Equivalent to the A1-A2 domain."
; NAME/KEY: Domain
; LOCATION: 1..2277
; OTHER INFORMATION: /note="CDNA encoding human factor
; OTHER INFORMATION: VIII."
US-08-474-503-1

Query Match 58.9%; Score 589; DB 1; Length 9009;
Best Local Similarity 96.0%; Pred. No. 1,1e-175;
Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 372 actgacatcaccttcttcttcttccacagtgatcatccatcgaatagagctc 431
DB 107 ATTAACCTTTGCTTCCAGTGAACATTTGATGACATAGTATGCAATAGAGCTCT 166
QY 432 ccactgtcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 491
DB 167 CCACCTGCTTTCTGCTGCTTGGGATTCCTTTAGGCCACCAAGAAAGTACTACC 226
QY 492 tgggtgcagtgaaatcgtatcagtgaaatcgtatcagtgaaatcgtatcagtg 551
DB 227 TGGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 286
QY 552 acgcaagattcttcttcttcttcttcttcttcttcttcttcttcttcttctt 611
DB 287 ACGCAAGATTCTCTCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 346
QY 612 aaagacatcttcttcttcttcttcttcttcttcttcttcttcttcttctt 671
DB 347 AAAAGACTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 406
QY 672 cctggtgagtgctgctgctgctgctgctgctgctgctgctgctgctgctg 731
DB 407 CCGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 466
QY 732 cacttaagaacatgctgctgctgctgctgctgctgctgctgctgctgctgct 791
DB 467 CACTTAAGAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 526
QY 792 aagcttcgagagagctgctgctgctgctgctgctgctgctgctgctgctgct 851
DB 527 AAGCTTCGAGGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 586
QY 852 aagcttcctggtgagagagagagagagagagagagagagagagagagagag 911
DB 587 AAGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 646
QY 912 tggctctgagagagagagagagagagagagagagagagagagagagagag 971

```

Db 647 TGGCCTTGACCCACTGTGCTTACCTACATCTTCTCATGTGAGACTGTGTAAG 706
 QY 972 actgaattcaggccctcatgtgagcccta 1000
 Db 707 ACTGAATTCAGGCTCATTTGAGACCCCTA 735

RESULT 13

US-08-670-707A-1
 ; Sequence 1, Application US/08670707A
 ; Patent No. 5839204
 ; GENERAL INFORMATION:
 ; APPLICANT: Lollar, John S.
 ; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/670,707A
 ; FILING DATE: 26-JUN-1996
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: NO PCT/US94/13200
 ; FILING DATE: 15-NOV-1994
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/212,133
 ; FILING DATE: 11-MAR-1994
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/864,004
 ; FILING DATE: 07-APR-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Greenlee, Lorraine L.
 ; REGISTRATION NUMBER: 27,894
 ; REFERENCE/DOCKET NUMBER: 75-95F
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 303/499-8080
 ; TELEFAX: 303/499-8089
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9009 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: cDNA to mRNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; TISSUE TYPE: Liver
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 5125..7053
 ; OTHER INFORMATION: /product= "Domain Structure"
 ; OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2 domain"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..2277
 ; OTHER INFORMATION: /product= "Domain Structure"
 ; OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..2277

OTHER INFORMATION: /product= "Domain"
 OTHER INFORMATION: /note= "cDNA encoding human factor-VIII"
 US-08-670-707A-1

Query Match 58.9%, Score 589, DB 2; Length 9009;
 Best Local Similarity 96.0%; Pred. No. 1,1e-115;
 Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 372 actgaatcaccactttcttcttccacaggtatcagatccaccatgaaatagactct 431
 Db 107 ATTAACTTTTGGCTTCTCCAGTGAACATTTGTACCAATAAGATGAATTAAGCTCT 166
 QY 432 ccaactgctcttcttggtcctttgcatctcgtttagtgcacccagaatactacc 491
 Db 167 CCACCTGCTTCTTCTGTGCGCTTTTGCGATCTCTGTTAGGCCACCAAGAAATACATACC 226
 QY 492 tgggtcagtggaactgcatgagactatagcaaaagtatcgtgltgagctgtg 551
 Db 227 TGGGTGAGTGAAGTGCATGATGAGGACATATATGCAAAAGTATCTGGTGAAGCTGCTG 286
 QY 552 acgcaagatttccctcctagatgccaataatcttccattcaacacctagctgtaca 611
 Db 287 ACGCAAGATTTCCTCCTAGAGTGGCCAAATCTTTCCATTCACACCTCACTGTTACA 346
 QY 612 aaaaagctcttcttgtagaattcaagacacacacacacacacacacacacacacac 671
 Db 347 AAAAGACTCTGTGTTAGTAATTCAGGTTACCTTTTCAACATGCTTAAGCCAAAGGCCAC 406
 QY 672 cctgagatggtctgtagtctcaccatccacagctgaggtttagatagacagtgatca 731
 Db 407 CCTGATGGGTCTGCTAGTCTCCTACCATCCAGGCTGAGGTTATATGATAGAGTGTACTTA 466
 QY 732 caactaaagaatggtctcccatcctcgtcaagcttcaatgctgtgtgtatccactcga 791
 Db 467 CACTTAAAGAAATGGCTTCCATCCATCTGTCAGTCTTCACTGTTGTTGTTATACCTACGTA 526
 QY 792 aagcttcgaggaagcctgaatatatgtagacacagctcaaaaggagaagaatagata 851
 Db 527 AAGCTTCGAGGAGACCTGAATATATGATGATGATGATGATGATGATGATGATGATGATG 586
 QY 852 aagcttcctcgtgtggaagcaccatacatatgtcgtgcaagctcgtgaaagaatgtc 911
 Db 587 AAGTCTCCCTGTTGAGGAGGACCATATATGATGATGATGATGATGATGATGATGATGATG 646
 QY 912 tggctctgaccacatggtgcttaccatcatatcttctcagtgagactgtgtaaa 971
 Db 647 TGGCCTTGACCCACTGTGCTTACCTACATCTTCTCATGTGAGACTGTGTAAG 706

RESULT 14

US-09-037-601-1
 ; Sequence 1, Application US/09037601
 ; Patent No. 6180371
 ; GENERAL INFORMATION:
 ; APPLICANT: Lollar, John S.
 ; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 ; STREET: 5370 Manhattan Circle Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9009 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: liver
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5125..7053
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2 domain"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2277
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2277
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "CDNA encoding human factorVIII"
US-09-037-601-1

```

	Query Match	58.9%	Score 589	DB 4	Length 9009
	Best Local Similarity	96.0%	Pred. No. 1.1e-175		
	Matches 604	Conservative	0	Mismatches 25	Indels 0
				Gaps	0
Qy	372	actacatccaccttcttcttcttcccaaggtatcgatccacatgcgaatagagctc	431		
Db	107	ATTACCTTTTGGTTCTCCAGTTGAACATTCTTACCAFAATGATCATGCAATATAGTCT	166		
Qy	432	ccacctgctcttcttctgaccttggacgtctgctttagtgccaacgaagatactacc	491		
Db	167	CCACCTCTCTCTTCTTGTCCTTTTGCGATTCTGCTTTAGTGCACCAAGAATACTACC	226		
Qy	492	tgggtggaagtgagacgtcatatggagactatatgcgaagatcctgggtgagctgctgttg	551		
Db	227	TGGTGTGAGTGAGACGTCTCATGGGACTATATGCAAAAGTATCTGGTGTAGCTGCTTGG	286		
Qy	552	acgcaagatttcttcccttagagatgcgaatcttcttccatccaaacctcagtcgtaca	611		
Db	287	ACGCAAGATTCTCTCCTAGAGTGCCAAACTTTTCATTCAACACCTCATGCTGTACA	346		
Qy	612	aaaagactctggtttagaattcaagatccgacttcttcaacatcgctgaagccaagccac	671		

Db	347	AAAAGACCTGTGTTGTGAATCAATCAAGGTTACCTTTTCAACATGGTAAAGCCAAAGGCCAC	406
Oy	672	ccctgatgagtcctgctctaagctctcaacatccaagctgagtttatgatacaagtygtcat	731
Db	407	CCGTGAATGGGCTGCTGTAAGTCCATCCACATCCAGGCTGAGTTAATGATACAGTGGTCTTTA	466
Oy	732	cacttaagaacaatgcttcccatccctgtaagcttcaagtcctgctgtgtggtatccctaagga	793
Db	467	CACTTTAAACAATGAGCTTCCATCCTGTGCACATCTTTCATAGCTGTGGTGTATTCCTACTGGA	526
Oy	792	aagctctcgaaggagctgaatatatgatcatcagaccagtcacaaggaggaagaagaatgata	851
Db	527	AAGCTTCGAGGAGAGCTAATATATATATGATGACAGCCAGTCAAAAGGAGGAAAGAAGATGATA	586
Oy	852	aagctctccctcgtgtggaagccatacatatgctctggcagctctcgtgaagaagaatgctcaa	911
Db	587	AAGTCTTCCCTGGGTGGAAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAGATGTCCAA	646
Oy	912	tggcctcagaccacatgctgaccttaacctactaatactcttctcattgtagacctgttaaag	971
Db	647	TGGCCTTCGACCACCATGTGCCCTTAACCTACTATATATCTTCTCATGTGAGACCTGGTAAAG	706
Oy	972	acttgaatcagacctcatatggaagcccta	1000
Db	707	ACTTGAATTCAGGCTCATTTGAGACCCCTA	735

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

TISSUE TYPE: Liver
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5001..7053
OTHER INFORMATION: /note- "Domain Structure:
OTHER INFORMATION: Equivalent to the A3-C1-C2 domain"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2277
OTHER INFORMATION: /note- "Domain Structure:
OTHER INFORMATION: Equivalent to the A1-A2 domain"
PCT-US93-03275-3

Query Match 58.9%; Score 589; DB 5; Length 9009;

Best Local Similarity 96.0%; Pred. No. 1.1e-175;

Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 372 actgacatccatttcttcttccacagatccatccatgcaatagagctct 431
DB 107 ATTAACCTTTGCTTCCAGTTGACATTTGACATATGATGCAATATGAGCTCT 166
QY 432 caacctgctcttctgctccttgcgaltcgtctagtcgaccagaagatactaco 491
DB 167 CCACCTGCTTCTTCTGCTTGGCTTTGGATTCGCTTTAGTGCCACAGAGATACTACC 226
QY 492 tgggtgagtgagacatgcatgagactatagcaagatgctcgtgagctgctgtg 551
DB 227 TGGGTGAGTGGAGACATGCTCATGAGACTATATGCAAGTATCTGGTGAAGCTGCTGTG 286
QY 552 acgcaagatttctcctcagagtgccaaatcttccattcaaacacctcagtcgtgtaca 611
DB 287 ACCGAGATTTCTCTAGAGTGCCTCAAAATCTTTCATTCACACCTCAGTGTGTACA 346
QY 612 aaaagactcgttltgtagaattcaagatcacttccaacatcgctaaagccaagycac 671
DB 347 AAAAGACTCTGTTGTAGAAATTCACGCTTCCACTTTCAACATCGCTAAAGCCAGGCCAC 406
QY 672 cctgagatgagtcgtcagagtcctacatccaggtcaggtttagatgataagtgatca 731
DB 407 CTGGAATGGGTCTGCTAGGTCTTACCATCCAGGCTGAGCTTATGATACAGTGTCTATTA 466
QY 732 caattaagaacatggttcccatcctcgtcagtcctcagtcctcgtgtgtatcctactga 791
DB 467 CACTTAAGAAACATGCGCTCCCATCTGTCAGTCTCATGCTGCTGTGCTATCTACTGGA 526
QY 792 aagctctgagagagctgataatgatgacagaccagtcacaagggagaagaagatgata 851
DB 527 AAGCTTCTGAGGAGGCTGAATATGATGATGACAGCAGTCAAGGGAGAAAGAGATGATA 586
QY 852 aagcttccctggtggaagccatataatgctcgtcaggtccctgaagaagaatgtgccaa 911
DB 587 AAGTCTTCCCTGCTGGAAGCATACATATGCTGGCAGGTCTCTGAAGAAGAAATGTGCCAA 646
QY 912 tggcctctgaccactgctccttaactactcalatcctcagtgagaccctgtaaaag 971
DB 647 TGGCCTCTGACCCACATGCTGCTTACTACTCAATCTTCTCATGTGAGACTGTGTAAGAAG 706
QY 972 acttgaattcaggtcctcattgagacccta 1000
DB 707 ACTTGAATTCAGGCCCTCATTTGGAGCCCTTA 735

Search completed: January 17, 2002, 16:27:40
Job time: 12370 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 14:58:51 ; Search time 3755.78 Seconds
(without alignments)
2861.131 Million cell updates/sec

Title: US-09-740-211-14_COPY_1_1000
Perfect score: 1000
Sequence: 1 cggccctgcagcagcagcgcg.....cagccctcatgtgagcccta 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estin: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_hic: *
10: gb_est1: *
11: gb_est2: *
12: gb_hic: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rtd: *
20: em_gss_vit: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	323.4	32.3	954	11	BG477974 602522966
2	98.8	9.9	456	10	AA817854 UI-R-A0-a
3	98.6	9.9	625	10	AT526753 u142e03.y
4	98.6	9.9	658	11	BG923886 602824585
5	98.6	9.9	783	11	BG977142 602846744
6	98.6	9.9	3739	12	AK014835 Mus muscu
7	98.2	9.8	1006	10	AT1225600 u10ec05.y
8	97.8	9.7	669	11	BF150467 uy86b07.y
9	97.7	9.6	632	11	BG862595 602795978
10	95.8	9.6	521	10	AA865338 CG88d10.s
11	95.8	9.6	878	11	BE912296 601666013
12	95.8	9.6	918	10	AL556703 AL556703

13	94.2	9.4	634	10	AV653260
14	93.8	9.4	477	10	AA461838
15	91.8	9.2	576	11	BE914540
16	91.6	9.2	913	11	BE916015
17	91.6	9.2	954	11	BG172588
18	90.4	9.0	722	11	BF136574
19	89.8	8.9	936	11	BF138391
20	89.2	8.9	695	11	BF107719
21	89.2	8.9	839	10	BE371524
22	89.2	8.9	836	11	BE917370
23	88.2	8.8	880	11	BF144665
24	88	8.8	1002	10	AA873581
25	88	8.8	657	11	BF532535
26	87.8	8.8	657	11	BF428548
27	87.6	8.7	621	11	BG914512
28	87.6	8.7	472	10	BE308899
29	86.6	8.7	530	10	BE285283
30	86.6	8.7	667	10	BF975622
31	86.6	8.7	691	11	BF975622
32	86.6	8.7	796	11	BF383789
33	86.6	8.7	837	11	BF144018
34	86.6	8.7	893	10	BE283339
35	86.2	8.6	843	11	BF158434
36	86	8.6	880	10	BE286923
37	86	8.6	732	11	BE568672
38	85.4	8.5	475	10	AT165337
39	85	8.5	670	11	BF976890
40	85	8.5	921	11	BF138986
41	84.6	8.5	921	11	BF156677
42	84.6	8.3	684	11	BF384924
43	83	8.3	509	11	BF604025
44	82.6	8.2	622	10	AW175422
45	81.6	8.2			

ALIGNMENTS

RESULT 1
LOCUS BG477974 954 bp mRNA EST
DEFINITION 602522966F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641352 5',
mRNA sequence.
ACCESSION BG477974 GI:13410253
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS 1 (bases 1 to 954)
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC/DCFTD/DPF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10K1408 row: c column: 17
High quality sequence stop: 626.
Location/Qualifiers
1. 954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4641352"
/clone_1ib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"

FEATURES

source

	DEFINITION						
		UJ42e03.y1 Sugano mouse liver mla Mus musculus cDNA clone IMAGE:1922620 5' similar to gb:M13699 CERUIOPLASMIN PRECURSOR (HUMAN)"; mRNA sequence.					
	ACCESSION	A1526753					
	VERSION	A1526753.1	GI:4440888				
	KEYWORDS						
	SOURCE	house mouse.					
	ORGANISM	Mus musculus					
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
	REFERENCE	1 (bases 1 to 625)					
	AUTHORS	Marras,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Thising,B., Allen,M., Bowers,Y., Peterson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.					
	TITLE	The WashU-NCI Mouse EST Project 1999					
	JOURNAL	Unpublished (1999)					
	COMMENT	Contact: Maria M./WashU-NCI Mouse EST Project 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouselst@wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:978912					
	FEATURES	Seq primer: custom primer used High quality sequence stop: 510.					
	SOURCE	Location/Qualifiers					
		1..625					
		/organism="Mus musculus"					
		/strain="C57BL"					
		/db_xref="taxon:10090"					
		/clone="IMAGE:1922620"					
		/clone_lib="Sugano mouse liver mla"					
		/sex="female"					
		/dev_stage="adult"					
		/lab_host="DH10B"					
		/note="Organ: Liver; Vector: pME18S-FL3; Site_1: DraIII (CACGTG); Site_2: DraIII (CACTATGG); 1st strand cDNA was primed with an oligo(dT) primer [ATGGCGCCATTTCCTTTTTTTT]; double-stranded cDNA was ligated into distinct draiii sites of the pME18S-FL3 vector (5' site CACTATGG, 3' site CACTATGG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCCTTAAGCTCG and 3' end primer CGACCCTGCACTCAGACA."					
	BASE COUNT	185 a 126 c 140 g 174 t					
	ORIGIN						
		Query Match 9.9%; Score 98.6; DB 10; Length 625; Best Local Similarity 55.0%; Pred. No. 8,3e-18;					
	Matches 216; Conservative 0; Mismatches 174; Indels 3; Gaps 1;						
Oy	608 tacaaaaagactctgttctgtagaatcaccgcacacctttccaatcgctaaggcaagg 667						
Dd	228 TATAAAGAAGCCCTTATTTTAGTGATGATGACAGATGCGACCTTAAAGACTTATAGACAAA 287						
Oy	668 caacctgatggtcgttgtagtccattaccatcacgaagtgcgaagtatatatacagtgtc 727						
Dd	288 CCAGCCCTGGCTAGGCTTTTAGGCCCTGTCACTCAAAGCTGAAGTTGAAGATTAAGTTTAT 347						
Oy	728 attacacttaagaacaatggtctcccactccttgtagctcatgctctgttgttatccctac 787						
Dd	348 GTTCACTTAAAGAACCTTGCTCTAGAGATCTACACTTCTCATGACATGGGGTAAAGCTAC 407						
Oy	788 ttgaaagctcttgaggagctgaatatgatgatcagacacgtaaccaaggagaaagaagt 847						

Db	Accession	Source	Organism	Reference Authors Title	Journal	Comment
Db	408	ACCAAGAGTATGAGGAGCGCTTACCTCGACAAACCACTGATTTTCAACGGGGTGAT	658 bp mRNA	EST	05-JUN-2001	
Qy	848	gataaagctctccctgttggaagccacacacatgctgtcgaagtcctgaaagatggt	mus musculus	CDNA clone	IMAGE:4953333	5'
Db	468	GACAAATGCTTCCCGGACAAACAGTATGTGTATGTCTGCTGCATGCC---AATGACCAAGT				
Qy	908	ccaatgctctgaagccacacatgctgacctaacacacatcttctatgtagtgcgtg				
Db	525	CTGTGGAAGAGGAGACACCAATTTGTGTATCCACGATTTTACACCTCCAGTGTGATCTCCA				
Qy	968	aaagactgtaattcaggccctcatggaagcccta	1000			
Db	585	AAAGATATTGCAATCAGACATCATAGGACCTCTA	617			
RESULT	4					
LOCUS	BG923886					
DEFINITION	602824585F1 NCL_CGAP_Mam6 Mus musculus					
ACCESSION	BG923886					
VERSION	1					
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
TITLE	1 (bases 1 to 658)					
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/.					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)					
	Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D.					
	Email: cgapb.r@mail.nih.gov					
	Tissue Procurement: Jeffrey Green M.D.					
	CDNA Library Preparation: Life Technologies, Inc.					
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)					
	DNA Sequencing by: Incyte Genomics, Inc.					
	Clone distribution: MGC clone distribution information can be					
	found through the I.M.A.G.E. Consortium/LNLN at:					
	http://image.lnl.gov					
	Plate: L1AM10912 row: j column: 22					
	High quality sequence stop: 657.					
FEATURES						
SOURCE						
	1..658					
	/organism="Mus musculus"					
	/strain="FVB/N"					
	/db_xref="taxon:10090"					
	/clone="IMAGE:4953333"					
	/clone_lib="NCI_CGAP_Mam6"					
	/sex="Female, Virgin"					
	/tissue_type="Infiltrating ductal carcinoma"					
	/dev_stage="5 months"					
	/lab_host="DH10B"					
	/note="Organ: mammary; Vector: PCMV-SPORE6; Site_1: SalI;					
	Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.					
	Library constructed by Life Technologies. Investigator					
	providing samples: Jeffrey Green, M.D., NIH"					
BASE COUNT	202 a 131 c 148 g 177 t					
ORIGIN						
Query Match	9.9%	Score 98.6:	DB 11:	Length 658:		
Best Local Similarity	55.0%:	Fred. No.8.5e-18:				
Matches 216:	Conservative 0:	Mismatches 174:	Indels 3:	Gaps 1:		
Qy	608	tacaaagacaccttctgtatgaattcaggaacaccccttccacacatgcgtacgaagc	667			
Db	228	TATTAAGAAAGCCCTTATTTTGTAGTACACAGATGGCACCTTTAGTAAGACTATPAGACAA	287			
Qy	668	ccacctgatggtctcgtagctccacacacacacacacacacacacacacacacacacac	727			
Db	288	CCAGCTCGGCTATGGGCTTTTATGAGCCCTGTCATCAACAGTGAAGATTGAAGATTAT	347			

Query Match	9.98;	Score 98.6;	DB 11;	Length 783;
Best Local Similarity	55.08;	Pred. No. 9e-18;		
Matches 216;	Conservative 0;	Mismatches 174;	Indels 3;	Gaps 1

MEDLINE
 20530913
 PUBMED
 11076861
 REFERENCE
 4 (bases 1 to 3739)

4 (bases 1 to 3739)
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.

DEFINITION AV653260 GLC Homo sapiens cDNA clone GLCDA09 3', mRNA sequence.
ACCESSION AV653260
VERSION AV653260.1 GI:9874274
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 634)
AUTHORS Qian,B., Wu,T., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,
Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,
Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
Homo sapiens cDNA clone
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
location/Qualifiers
1. 634
FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCDA09"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 213 a 119 c 128 g 172 t 2 others
ORIGIN
Query Match 9.4%; Score 94.2; DB 10; Length 634;
Best Local Similarity 52.4%; Pred. No. 1.6e-16;
Matches 207; Conservative 0; Mismatches 188; Indels 0; Gaps 0;
QY 606 tttcaaaaagactctgtttagaattcagatccagtcctttcaaacatgcgtaagcaa 665
DB 203 TATATAAGAAAGCCCTTATCTTCAGTACACAGATGAACCTTTAGACAACTATTAGAAA 262
QY 666 ggcacacctgagatggtctgtaagctcctacatccagctgaggtttatgatacagtg 725
DB 263 AACCGGCTGGCTTGGGTTTATAGCCCTATATCAAAAGCTGAAGAACTGGAGATTAAGTTT 322
QY 726 tcattacacttaagaacatggtctccatctgtcagctcttcacatgctggtggtatcct 785
DB 323 ATGTACACTTAAAAAACCCTTGCTAGGCCCTTACACCTTTCAATTCACATGGAATTAACCT 382
QY 786 actgaagaagctctcgagagagctgaataatgatacagaccagcagaaggaagaagaag 845
DB 383 ACTATAAGAGACATGAGAGGGCCATCTACCTGTATACACACAGACATTTTCAAGAGACAG 442
QY 846 atgataaagcttcctctggtgtaagcacaataatgtctgagcaggtctgtaagaagatg 905
DB 443 ATGACAAAGTATTTACAGAGAGAGAGATATACATGATGTTGCTTGCACAGAGAAACACA 502
QY 906 gtccaatgagctctgagcccaactgctcctacatctcatalctttccacatgtagactgt 965
DB 503 GTTCGGGGAAGAGATGCAATTGTGACTAGATTTACCAATTTCCACATGATGATGCTC 562
QY 966 taaaagacttaattcaagcctcatgtgagcccta 1000
DB 563 CAAAGATATTTGCTCAGAGACTCAATCGACCTTTA 597

RESULT 14
AA461838

LOCUS AA461838 477 bp mRNA EST 10-JUN-1997
DEFINITION v95d04.r1 Soares_mammary_gland_NbMNG Mus musculus cDNA clone
IMAGE:851527 5', similar to gb:U13699 CERUOLOPLASMIN PRECURSOR (HUMAN
);, mRNA sequence.
ACCESSION AA461838
VERSION AA461838.1 GI:2186729
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 477)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,B.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wylie,T., Lennon,G., Soares,B., Willson,R. and
Waterston,R.
The WashU-HMMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 850L, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (Info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 413.
location/Qualifiers
1. 477
FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:851527"
/clone_lib="Soares_mammary_gland_NbMNG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DHI0B"
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer 15.
TGTATCAATCTGAAGTGGAGGCGCCCGCAATGTTTGTGTTTGTGTTTGTGTTT
T 371; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."

BASE COUNT 148 a 102 c 102 g 125 t
ORIGIN
Query Match 9.4%; Score 93.8; DB 10; Length 477;
Best Local Similarity 54.5%; Pred. No. 1.9e-16;
Matches 210; Conservative 0; Mismatches 172; Indels 3; Gaps 1;
QY 608 tacaanaagactctgtttagaattcagatccacatttcaacatgcgtaagccaagg 667
DB 77 TATTAAGAGGCCCTTATTTTGGTAGACAGATGAGCACTTTAGTAGACTATAGACAAA 136
QY 668 ccacacctgagatggtctgtagctcctaccatccagctgaggtttatgatacagtgatc 727
DB 137 CCAGCCTGGCTAGGTTTATTTAGCCCTGTCATCAAGCTGAAGTTGAATTAAGTTTAT 196
QY 728 attacacttaagaacatggtctccatcctgctgtagctcatgctggtgtagtatacctac 787
DB 197 GTTCACTTAAGAGACCTTGCTCTAGAGATCTACACTTTTCAATGACATGGGATACGTAC 256

Oy 788 tgyaaattcttbaaggaagctaaatctagtgtctgaccagctcaaggagaaagaat 847
 |||||
 Db 257 ACCAAGAGATATAGGAGGAGCCGCTTACCTGCAACACCAGCTGATTTTCAACGGGCTGAT 316
 Oy 848 gataaagcttccctctgttgyaaagcatacatatgtctgtgcaggtccctgaagaaatggt 907
 |||||
 Db 317 GACAAATGCTTCCCGGACACAGTATGTATGTGTGTGTCAGGCC---AATGAGCCAGT 373
 Oy 908 ccaatgtccctctacccaactgtgtccttaactcctatatcttctcatgtgagacctgta 967
 |||||
 Db 374 CCTGGAAGGGAACAGCAATTTGTGTGACCAAGATTTACCATCCCATGTTGATGTCACA 433
 Oy 968 aaagacttgaattcgaagcctcatgt 992
 |||||
 Db 434 AAGATATTTCATCAGGACTCATAG 458

RESULT 15	BE914540	LOCUS	DEFINITION
	BE914540	576 bp	mRNA
	601665658f1	NCI_CGAP_Mam1	Mus musculus cdna clone IMAGE:3965721 5'
	mRNA sequence.		

VERSION	BE914540.1	GI:104132700
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus
AUTHORS	1 (Pages 1 to 576)
TITLE	NIH-MGC http://mgc.nhl.nih.gov/
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg Ph.D.

```

FEATURES
source
tissue Procurement: Gilbert Smith, Ph.D.
cdna Library Preparation: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LMAG137 row: 1 column: 10
High quality sequence stop: 576.
Location/Qualifiers
1..576

```

```

/organism="Mus musculus"
/strain="FVB/N"
/db xref="taxon:10090"
/clone="IMAGE:3965721"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DH10B"
/name="Organ: mammary; Vector: pCMV-Sport6; Site:1: Saliv
Site:2: Notti; Cloned unidirectionally. Primer: oligo dr
library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      178 a      119 c      127 g      151 t      1 others
ORIGIN

```

Query Match 9.2%; Score 91.8; DB 11; Length 576;
Best Local Similarity 53.9%; Pred. No. 7.9e-16;
Matches 212; Conservative 0; Mismatches 177; Indels 4; Gaps 1;

QY	Db	QY	Db
608	668	157	217
tacaaaaagcctcgttllttagataltcaagatcaccccttcaacatgcgtacgcaag	cgacccctgtagtgcctgcgaagtcctacacatccagcgtgaagttatgatcaagtgtgc	TATAGAGAGCCCTTTATTTTGGATGACACAGATAGGCACCTTTGTTAAGACTATAGACAAA	CCAGCCCTGGCTAGGAGTTTATTTAGGCCCTGCTATCTTAAAGCTGAACTTGAAGGTTTAT
667	727	216	276

OY	728	atfacacttaagaacacatgagctctcccatccctgtaagctctcaagctgtgttggtatccctac	787
Db	277	gttcactttaaagaacaccttgcctctgagatcttaccacttttttcaatgcacatggggatrac	336
OY	788	tgsaaagctctctgsgaagctgaatatgatgatcagaacagatcaaaagsgaaagaat	847
Db	337	accgaaggatattgagggagccgtctaccctgcacacacccactgatttttccacaggcgtat	396
OY	848	gataaagcttcctccgtgtgtgaagcacaataatgtctcgaagcttcctgaagaagaatgct	907
Db	397	gacaaagctgtctcccggaacacataatgtgtatgtgctgcataagccaatgaccagctct	456
OY	908	ccaattgagctcttgacccaactgtgccttaacbaactacataatcttctcaatgttgagcctgta	967
Db	457	ggaagagc---GAAcAGcAGcAATgTGtATcACCAGAGATTTCACCACTCCATGTTGATGCTCA	512
OY	968	aaagactgaatcagagccatcttgagaccccta	1000
Db	513	AAAGATATGTGATCAGGACCTATAGGACCTCTA	545

```
Search completed: January 17, 2002, 14:58:55
Job time: 8656 sec
```

Job time: 8656 sec

Tue Jan 22 15:12:43 2002

us-09-740-211-14_copy_1_1000.rst

Page 11

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:26:26 ; Search time 5101.32 Seconds
(Without alignments)
3233.904 Million cell updates/sec

Title: US-09-740-211-14_COPY_4000_4999
Perfect score: 1000
Sequence: 1 gacacagtgatcgtctgtg.....cgcagctgctgcagcagacat 1000

Scoring table: IDENTITY-NTC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: GenBank:
2: gb_ba:
3: gb_hlg:
4: gb_in:
5: gb_cm:
6: gb_ov:
7: gb_pat:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_om:
20: em_or:
21: em_ov:
22: em_pat:
23: em_ph:
24: em_pl:
25: em_ro:
26: em_sts:
27: em_sy:
28: em_un:
29: em_vl:
30: em_hlgo_hum:
31: em_hlgo_inv:
32: em_hlgo_rnd:
33: em_hlg_hum:
34: em_hlg_inv:
35: em_hlg_rnd:
36: em_hlg_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1000	100.0	4999	6	ARI38378	ARI38378 Sequence
2	1000	100.0	4999	6	ARI46888	ARI46888 Sequence
3	954.4	95.4	11933	6	ARI38377	ARI38377 Sequence
4	954.4	95.4	11933	6	ARI46887	ARI46887 Sequence
5	790	79.0	4670	6	ARI10040	ARI10040 Sequence
6	790	79.0	5035	6	AR034084	AR034084 Sequence
7	790	79.0	5035	6	AR071306	AR071306 Sequence
8	790	79.0	5035	6	ARI12722	ARI12722 Sequence
9	790	79.0	6999	6	I77105	I77105 Sequence 1
10	790	79.0	7056	6	I27063	I27063 Sequence 1
11	790	79.0	7272	6	I05404	I05404 Sequence 3
12	790	79.0	8967	6	AX052730	AX052730 Sequence
13	790	79.0	8967	6	I71409	I71409 Sequence 1
14	790	79.0	8967	9	HSEFVILIR	X01179 Human mRNA
15	790	79.0	9009	6	AR003710	AR003710 Sequence
16	790	79.0	9009	6	AR029065	AR029065 Sequence
17	790	79.0	9009	6	ARI26884	ARI26884 Sequence
18	790	79.0	9009	6	I31901	I31901 Sequence 3
19	790	79.0	9009	6	I63424	I63424 Sequence 1
20	790	79.0	9009	9	HUMFVIIIC	K01740 Human coagu
21	790	79.0	9029	9	HUMFVIIIC	M14113 Human coagu
22	790	79.0	9354	6	AR003585	AR003585 Sequence
23	788.4	78.8	8241	6	A05328	A05328 Synthetic F
24	788.4	78.8	8241	6	A07042	A07042 Artificial
25	787	78.7	4278	6	I08644	I08644 Sequence 4
26	787	78.7	4551	6	I08641	I08345 Sequence 1
27	786.8	78.7	7440	6	I08345	I08345 Sequence 2
28	785.4	78.5	4281	6	I08643	I08643 Sequence 3
29	785.4	78.5	4548	6	I08642	I08642 Sequence 2
30	783.6	78.4	7440	6	I08457	I08457 Sequence 1
31	695.4	69.5	1728	6	I02054	I02054 Sequence 8
32	634.8	63.5	7493	6	AR003712	AR003712 Sequence
33	634.8	63.5	7493	6	AR029067	AR029067 Sequence
34	634.8	63.5	7493	6	ARI26886	ARI26886 Sequence
35	634.8	63.5	7493	6	I63427	I63427 Sequence 7
36	634.8	63.5	7493	10	MUSCIVIII	I05573 Mus domesti
37	628.4	62.8	7493	9	HUMF8C	M90707 Homo sapien
38	619.8	62.0	7145	4	AF049489	AF049489 Canis fam
39	614	61.4	7032	4	AF016234	AF016234 Canis fam
40	594.8	59.5	4334	6	AR029098	AR029098 Sequence
41	594.8	59.5	4334	6	ARI26917	ARI26917 Sequence
42	594.8	59.5	6402	6	AR029097	AR029097 Sequence
43	594.8	59.5	6402	6	ARI26916	ARI26916 Sequence
44	594.8	59.5	6539	4	SS049517	U49517 Sus scrofa
45	592	59.2	8831	6	E00527	E00527 Complete cd

ALIGNMENTS

RESULT 1
LOCUS ARI38378 4999 bp DNA
DEFINITION Sequence 14 from patent US 6200560.
ACCESSION ARI38378
VERSION ARI38378.1 GI:14480723
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4999)
Couto,L.B., Colosi,P.C. and Qian,X.
TITLE Adeno-associated virus vectors for expression of factor VIII by target cells
JOURNAL Patent: US 6200560-A 14 13-MAR-2001;
FEATURES
source location/Qualifiers
BASE COUNT 1375 a 1151 c 1155 g 1318 t
ORIGIN

Query Match 100.0%; Score 1000; DB 6; Length 4999;
 Best Local Similarity 100.0%; Pred. No. 1.2e-299;
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gataaagtgatctgttgacccaatgattatccagcgatcaagaagtgccg 60
 DB 4000 GATCAAGGTGATCTGTGGCACCATTATTCACGGGATCAACACCGAGGTCCCG 4059

QY 61 tcaagaattccagcctcaacatctcctcaagttatcatcatgatagtatggaa 120
 DB 4060 TCAGAAATTCAGGCTTACATCTCTCAAGTTATCATCATGATATAGTCTGATGGAA 4119

QY 121 gaagtgagacactatcgagaagaattccactgaaccttaagtgttcttggaat 180
 DB 4120 GAAGTGGACACTTATCGAGAAATTCACAGAACCTTAATGCTTTCTTGACATGT 4179

QY 181 ggaatcatctgggaataaacaataattttaacctccaattatgtctcgaatac 240
 DB 4180 GGATTCATCTGGGATTAACACAAATTTTAACTCCCAATATATGCTGATACATCG 4239

QY 241 ttggacccaactcatatagatctggacgaactcttgcaaggagtgaaggctgga 300
 DB 4240 TTGGACCCCACTCATTAATGATGCGACACTCTTCCATGGAGTTGATGGCTGTGA 4299

QY 301 tttaaatgttcagagatgcatctggaaatgagaataagaataatcatcgatcga 360
 DB 4300 TTAAATAGTTCAGCATGCCATTTGGGAATGGAGTAAGCAATATCATGATGCAAGAT 4359

QY 361 tactgttatacttaacttaacaatatgttgcacactgtgtctcctcaaaagtcga 420
 DB 4360 TACTGTATCTCTTACTTACCAATATGTTTGCACACTGCTCTCCTCAAAAGCTGCACT 4419

QY 421 tcaacccaagaagagatgaatgctggagacactcagatgaataatcaaaagatgag 480
 DB 4420 TCACCTCCAAAGGAGAGAGATATGCTGGAGACCTCAGATGAATATCCAAAGAGTGGCT 4479

QY 481 gaaagtgaacttcagaaagaacaatgaatgcaagagtaactcagaggaatgaatc 540
 DB 4480 GCAAGTGGACTTCCAAAGACAAATGAAGTCAAGAGTAATCTCAGGGAGTAAATTC 4539

QY 541 ttgtgttaccacatgatagtgaagagttcctcatctccagcagtcgaagtgcacat 600
 DB 4540 TGTGCTTACACATGATGTAAGAGAGTCTCTCATCTCCAGCACTCAAGATGGCATCA 4599

QY 601 gtagactctcttttccagaatgcaagaatgaagtttccagggaaatcaagactcct 660
 DB 4600 GTGAGACTCTCTTTTTCAGAAATGCAAAAGTAAAGTTCAGGGAATTCAGACTCTCT 4659

QY 661 caacactgtgtgaactctctagaccacgltactgaactgactcgaattcacc 720
 DB 4660 CACACCTGTGGAGAACTCTCTAGACCCGTTACTGACTGCTTCAATTCACCC 4719

QY 721 ccaagattggatgacacagattgcccgaagatgagtttggctgctggaggaagga 780
 DB 4720 CCAGAGTTGGGTGACCAATTTGCCCTGAGATGAGGTTTGGCTGCGAGCAAGGA 4779

QY 781 ccttactgaactcgaacttaataaagaatatttcatctgaatagttgtgtgtt 840
 DB 4780 CCTTACTGACTCGAGCTTAATTAAGAAATTTTATTTTCATTTGCAATATGTTGTTG 4839

QY 841 ttgtgtgctggcgcgaagaacccctagtatgtagagttggcactccctctctgcgc 900
 DB 4840 TTTGTGTGCTGGCCTGAGAACCCCTAGATGAGTTGGCCACTCCCTCTCTGCGCGCTC 4899

QY 901 gctcgtcactcgaagcgcgcgaaccaagatcgccgcgaagcgcgcgcgcgcgcgc 960
 DB 4900 GCTCCTTACTGAGGCTGGGCGACCAAAAGTCCGCCGCGCGCGGCTTGGCGGCGG 4959

QY 961 cctcaagtgagcgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1000
 DB 4960 CCTCAGTGAAGGAGGAGCGGCGGAGCTGCTCGAGGACAT 4999

RESULT 2
 ARI46888
 LOCUS ARI46888 4999 bp DNA PAT 08-AUG-2001
 DEFINITION Sequence 14 from patent US 6221349.
 ACCSSION ARI46888
 VERSION ARI46888.1 GI:15110691
 KEYWORDS
 SOURCE
 ORGANISM Unknown.
 REFERENCE
 AUTHORS 1 (bases 1 to 4999)
 TITLE Couto, L.B., Colosi, P.C. and Qian, X.
 JOURNAL Adeno-associated vectors for expression of factor VIII by target
 FEATURES
 source location/Qualifiers
 BASE COUNT 1375 a 1151 c 1155 g 1318 t
 ORIGIN

Query Match 100.0%; Score 1000; DB 6; Length 4999;
 Best Local Similarity 100.0%; Pred. No. 1.2e-299;
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gataaagtgatctgttgacccaatgattatccagcgatcaagaagtgccg 60
 DB 4000 GATCAAGGTGATCTGTGGCACCATTATTCACGGGATCAACACCGAGGTCCCG 4059

QY 61 tcaagaattccagcctcaacatctcctcaagttatcatcatgatagtatggaa 120
 DB 4060 TCAGAAATTCAGGCTTACATCTCTCAAGTTATCATCATGATATAGTCTGATGGAA 4119

QY 121 gaagtgagacactatcgagaagaattccactgaaccttaagtgttcttggaat 180
 DB 4120 GAAGTGGACACTTATCGAGAAATTCACAGAACCTTAATGCTTTCTTGACATGT 4179

QY 181 ggaatcatctgggaataaacaataattttaacctccaattatgtctcgaatac 240
 DB 4180 GGATTCATCTGGGATTAACACAAATTTTAACTCCCAATATATGCTGATACATCG 4239

QY 241 ttggacccaactcatatagatctggacgaactcttgcaaggagtgaaggctgga 300
 DB 4240 TTGGACCCCACTCATTAATGATGCGACACTCTTCCATGGAGTTGATGGCTGTGA 4299

QY 301 tttaaatgttcagagatgcatctggaaatgagaataagaataatcatcgatcga 360
 DB 4300 TTAAATAGTTCAGCATGCCATTTGGGAATGGAGTAAGCAATATCATGATGCAAGAT 4359

QY 361 tactgttatacttaacttaacaatatgttgcacactgtgtctcctcaaaagtcga 420
 DB 4360 TACTGTATCTCTTACTTACCAATATGTTTGCACACTGCTCTCCTCAAAAGCTGCACT 4419

QY 421 tcaactccaagagagagatgaatgctggagacactcagtggaataatccaaaagatgag 480
 DB 4420 TCACCTCCAAAGGAGAGAGATATGCTGGAGACCTCAGATGAATATCCAAAGAGTGGCT 4479

QY 481 gaaagtgaacttcagaaagaacaatgaatgcaagagtaactcagaggaatgaatc 540
 DB 4480 GCAAGTGGACTTCCAAAGACAAATGAAGTCAAGAGTAATCTCAGGGAGTAAATTC 4539

QY 541 ttgtgttaccacatgatagtgaagagttcctcatctccagcagtcgaagtgcacat 600
 DB 4540 TGTGCTTACACATGATGTAAGAGAGTCTCTCATCTCCAGCACTCAAGATGGCATCA 4599

QY 601 gtagactctcttttccagaatgcaagaatgaagtttccagggaaatcaagactcct 660
 DB 4600 GTGAGACTCTCTTTTTCAGAAATGCAAAAGTAAAGTTCAGGGAATTCAGACTCTCT 4659

QY 661 caacactgtgtgaactctctagaccacgltactgaactgactcgaattcacc 720

Db	Accession	Sequence	Length	Source
Db	4660	CACACCTGTGTGTAACCTCTCTAGAACCCACCGTTACTGTACTGTACCTACCTTCGAATTTACACC	4719	GenBank
Qy	721	cccaagattgggtgtcaccagaatttgcctctggaagatgtgaagttcttgggtcttgcagagacaga	780	GenBank
Db	4720	CCAAAGTTGGGTGTGACCAACGATTTGCCCTGAGGATGAGGATTTGGGACTGCGAGGCACAGGA	4779	GenBank
Qy	781	cctctactgactcgcagccctaaataaaggaatatlttcatlgtcaatagtgltgltt	840	GenBank
Db	4780	CCCTCTACTGACTCGACGCTATTAAAGGAAATTTATTTTCATTGCAATAGTGTGTGGCTTT	4839	GenBank
Qy	841	tttgtgtgctgcgcagagaaacccctctagtgtgagtgttgccacccctctctgagcgtc	900	GenBank
Db	4840	TTTTGTGTGCTGGCGCCACGAGAACCCCTAGTGAATGAGATTGGCACTCCCTCTCTGTGGCGGCTC	4899	GenBank
Qy	901	gctcgctacactgagccgcagcgagaccaaagtgctgcgcgcagccgcgcgcttgcgccgagcgcg	960	GenBank
Db	4900	GCTGCGCTACACGAGGCGCGGCGGACCCAAAGGTGTGCCCGCAGCGCCGGGCTTTGCGCGGGCGG	4959	GenBank
Qy	961	cctcaagtgagcgagcgagcgagcgagcttgcctctgcagagcat	1000	GenBank
Db	4960	CCTCAGTGAGCGAGCGAGCGAGCGGACACTCTCTCAGACAGACAT	4999	GenBank
RESULT	3			
LOCUS	AR138377			
DEFINITION	AR138377	Sequence 13 from patent US 6200560.		
ACCESSION	AR138377			
VERSION	AR138377.1	GI:14480722		
KEYWORDS				
SOURCE				
ORGANISM		Unknown.		
		Unclassified.		
REFERENCE		1 (bases 1 to 11933)		
AUTHORS		Couto, L.B., Colosi, P.C. and Qian, X.		
TITLE		Adeno-associated virus vectors for expression of factor VIII by target cells		
JOURNAL		Patent: US 6200560-A 13 13-MAR-2001;		
FEATURES		Location/Qualifiers		
		1..11933		
		/organism="unknown"		
BASE COUNT		3258 a 2818 c 2717 g 3140 t		
ORIGIN				

Query Match	95.4%	Score 954.4	DB 6	Length 11933
Best Local Similarity	97.9%	Prod. No. 2.2e-285		
Matches	979	Conservative	0	Mismatches 16; Indels 5; Gaps 1
QY	1 gatcaagtgatgctctgttgcacacatgatatltcaagcgatcaccaagcaggtgccg	60		
DB	3975 GATCAAGTGGATCTGTTGGCACCAATGATTTTCACGGCATCAAGACCCAGGTTCCCG	4034		
QY	61 tcagaagttctccagcctctacatctctcagttatcaltcatgatatgcttgaaggaa	120		
DB	4035 TCAGAAGTCTTCACGGCTTACATCTCTCAGTTTATCATCAATGATAGCTTGATGGAA	4094		
QY	121 gaagtgacagatatttcgggaattccacgtgaaccttaatggtctctttggcaatt	180		
DB	4095 GAAGTGGCACTTATTCGAGGAATTCACCTGGAACCTTAATGATCTCTTTGGCAATGT	4154		
QY	181 gattcatcttgataataaacacatatlttlaaaccttccaatatattgctcgatataccg	240		
DB	4155 GATTCTATCTGGATTAACACATAATTTTAAACCTCCAAATTATGTGCGATACATCCG	4214		
QY	241 ttctgacccaactcaltatagcatlctgcagcactcttcgcatvgagttgatggctgtga	300		
DB	4215 TTTCACCCCAACTATTATAGCATTTTCGACGACACTCTTCGATGAGATGTGATGGCTGTGA	4274		
QY	301 tttaaatagttgagatgcacatttggaattggagatgaagcaatctcgatcacagat	360		
DB	4275 TTTTAAATAGTTGGAGATGCGATTGGGAATGGAGATGAAGCAATATCAATGACACAGAT	4334		

OY	361	tactgtcttaaccttaactttcccaatatggtttgcacacttgcttccttcaaaagctgcact	420
Db	4335	TACTGCTTCACTCACTTACCAATAATGTTGGCCACTGGTCTCTTCAAAAGCTGCAC	439
OY	421	tcacctccaagaaggaggaatagtccttgtagaacctcaggtgtaataatccaagaagtgct	480
Db	4395	TCACCTCCAAAGGAGGAGTAATGCTCTGGAGACCTCAGTGAAATATCAAAAGATGGCT	445
OY	481	gcaagttgacttcagaaagacataagaagttcacaggggtactactaagtaggaataatc	540
Db	4455	GCAAGTTGACTTCCAGAAAGACAAAGAAAGTCAAGGAGTAATCACTCAGGAGTAATC	451
OY	541	tctgtttcagaactgtatgtgtgaaggagttcccatctcagcagtcagtaagaatggtcaatc	600
Db	4515	TCTGCTTACCAACATATATGTGAAGAGATTCTCATCTCCAGCAGTGAAGTGGCAATCA	457
OY	601	gtgagctctctttttccagaatggcaaaagtaaaagtttttcagggaatacaagaactcct	660
Db	4575	GTGAGCTCTCTTTTTCAGAAATGGCAAAAGTTAAAGTTTTCAGGGAATCAAGACTCTCTT	463
OY	661	caacaccttggtgaactctctcagaagcccaacggttaactgaactgcacttcgaattcaacc	720
Db	4635	CACACCTGTGGGAAGACTCTTAGAACCCACCTTACTGACTGCTACTCTTGAATTACCC	469
OY	721	ccaagatttggtgtgcaccagaattgcctctgaatgtagatgtagatcttgggtctggaagcaaga	780
Db	4695	CCAGAGTTGGGTGCACCAAGATGGCCCTGAGATGGAGTGTCTGGGGCTGCGAGCACAGGA	475
OY	781	cctctactgactcggagcctataaaggaatttttcatctatgtcaatagtggtgtgttt	840
Db	4755	CCCTCACTGACTCGAGATATAAAGTTCAGACTCTAGAGAT-----CTGTGTGTGGTTT	481
OY	841	tttgtgtgagcgcgaagaaacccctagtagatgtagatgttgagcaactccctctctgcgctc	900
Db	4810	TTTGTGTGGCGCGGAGGAACCCCTAAGTAGATGAGTTGGCACATCCCTCTGCGCGCTC	486
OY	901	gctcgctcaactgaagcgcgggcgagaccaaaagtcgcgcgaagcccggtcttgcgcgggcg	960
Db	4870	GCCTCGCTCACTAGAGCGCGGCGCACCAAAAGTGTGCCCGCAGCCCGGGCTTTGTCCCGGGGG	492
OY	961	cctcattgagcagcagcgcgcgcagctccttgcagagaact 1000	
Db	4930	CCCTCACTGAGCAGCGCAGCGCCGACCTGCCTCTCAGAGAAAT 4969	

[illegible]

Db 4433 CCTACTGAGGTTGGCC 4450

RESULT 6

LOCUS AR034084 5035 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5869292.
ACCESSION AR034084
VERSION AR034084.1 GI:5949689
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Voorberg, J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;
FEATURES Location/Qualifiers
Source 1..5035
/organism="unknown"

BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2,7e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagttgatctgtttgacccaatgattatcagcgacacagccagggtgccc 60
DB 4231 GATCAAGTTGATCTGTTGGACCAATGATTATTCAGGCAATCAAGCCAGGTTGCCG 4290
QY 61 tcagaagttctccagccctacatctcagttatcaatcagttatgatttcttgatgaa 120
DB 4291 TCAGAAGTTCTCCAGCCCTACATCTCAGTTATTCATGATGATGATGATGAA 4350
QY 121 gaagtggcagactatcagaggaatccacatggaacctaatgctctcttggaatgt 180
DB 4351 GAAGTGGCAGACTATCAGAGGAATTCACATGCAACCTTAATGCTCTTGGCAATGT 4410
QY 181 ggattcatctggatgaataacaacatattttaaccctcaattatctgcatacatcg 240
DB 4411 GGATTTCATCTGGGATTAACAATATTTTAACTCCCAATATTTGCTCATATATCG 4470
QY 241 ttggacccaactcatatagcattcgacgacatcttcgcatgtagttgagtgctgtga 300
DB 4471 TTGGACCCAACCTCATATAGCATTCGACGACATCTTCGATGAGTGTGATGGCTGTGA 4530
QY 301 ttttaaatgttcagcagatgccattgggaatggagtaagaacatatacagatgacagat 360
DB 4531 TTTAAATGTTGCAGCATGCCATTTGGGAATGGAGTAAGCAATATCAGATGCAAGAT 4590
QY 361 tactgttcatctcactattacccaatatgtttgcacactggtctcctcttaaaagctgact 420
DB 4591 TACTGTTTCATCTCTATTACCAATATGTTTGCACCTGCTCTCTCTCAAAAGCTGCAGT 4650
QY 421 tcaactcgaaggaggaatgactgttcgagacccaagtgatataatccaaaagatgct 480
DB 4651 TCACCTCGAAGGAGAGTAATGCTCGAGACCTCAGGTAATATCCAAAAGATGGCT 4710
QY 481 gcaagtgagactccagaagaacaatgaagtcacagagtaactcagaggtgtaaatc 540
DB 4711 GCAAGTGAGACTCCAGAAACAATGAATTCACAGAGTAATCTACTCAGGAGTAATATC 4770
QY 541 tctgcttacacagatglatgtgaaggatctccatctccagcaagtcagaatgagcatca 600
DB 4771 TCTGCTTACACAGATGTATGTGAAGAGTTCCTCATCTCTCAGCAGATCAAGATGGCATCA 4830
QY 601 gtgagactctcttttccaaatgagcaaatgaagttttcaggggaaatcaagactccct 660
DB 4831 GTGAGACTCTCTTTTCCAAATGAGCAAGTAAGGTTTTCAGGGAATATCAAGACTCTT 4890

QY 661 cacactgtgtgtaactctcttagaccacacgttactgactcgtactcgaattacccc 720
DB 4891 CACACTGTGTGTAACCTCTCTTAGACCACACCGTTACTGACTCGTACTCGAATTCACCC 4950

QY 721 ccagaagttgggtgacacagattgcccctgagatgagtgagttcttggtctcgagggcacagga 780
DB 4951 CCAGAAGTTGGGTGACACAGATTGCCCTGAGAGTAGAGTGTGCTCGAGGACACAGGA 5010

QY 781 cctctactgactgagcc 798
DB 5011 CCTCTACTGAGGTTGGCC 5028

RESULT 7

LOCUS AR071306 5035 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5910481.
ACCESSION AR071306
VERSION AR071306.1 GI:7222194
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Voorberg, J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5910481-A 1 08-JUN-1999;
FEATURES Location/Qualifiers
Source 1..5035
/organism="unknown"

BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2,7e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagttgatctgtttgacccaatgattatcagcgacacagccagggtgccc 60
DB 4231 GATCAAGTTGATCTGTTGGACCAATGATTATTCAGGCAATCAAGCCAGGTTGCCG 4290
QY 61 tcagaagttctccagccctacatctcagttatcaatcagttatgatttcttgatgaa 120
DB 4291 TCAGAAGTTCTCCAGCCCTACATCTCAGTTATTCATGATGATGATGATGAA 4350
QY 121 gaagtggcagactatcagaggaatccacatggaacctaatgctctcttggaatgt 180
DB 4351 GAAGTGGCAGACTATCAGAGGAATTCACATGCAACCTTAATGCTCTTGGCAATGT 4410
QY 181 ggattcatctggatgaataacaacatattttaaccctcaattatctgcatacatcg 240
DB 4411 GGATTTCATCTGGGATTAACAATATTTTAACTCCCAATATTTGCTCATATATCG 4470
QY 241 ttggacccaactcatatagcattcgacgacatcttcgcatgtagttgagtgctgtga 300
DB 4471 TTGGACCCAACCTCATATAGCATTCGACGACATCTTCGATGAGTGTGATGGCTGTGA 4530
QY 301 ttttaaatgttcagcagatgccattgggaatggagtaagaacatatacagatgacagat 360
DB 4531 TTTAAATGTTGCAGCATGCCATTTGGGAATGGAGTAAGCAATATCAGATGCAAGAT 4590
QY 361 tactgttcatctcactattacccaatatgtttgcacactggtctcctcttaaaagctgact 420
DB 4591 TACTGTTTCATCTCTATTACCAATATGTTTGCACCTGCTCTCTCTCAAAAGCTGCAGT 4650
QY 421 tcaactcgaaggaggaatgactgttcgagacccaagtgatataatccaaaagatgct 480
DB 4651 TCACCTCGAAGGAGAGTAATGCTCGAGACCTCAGGTAATATCCAAAAGATGGCT 4710
QY 481 gcaagtgagactccagaagaacaatgaagtcacagagtaactcagaggtgtaaatc 540
DB 4711 GCAAGTGAGACTTCAGAAACAATGAATGAAGTACAGAGTAATCTACTCAGGAGTAATATC 4770

QY 541 tctgtaccagcagatgataatgaagaggttcctcaatcctccagcagtaagaatgacata 600
|||||
Db 4771 TCTGCTTACCAAGCATATATGTAAGAGGTTCCTCATCTCCAGCAGTCAGATGSCATCA 4830
QY 601 gtgagctctcttttttcaagaatgcaagtaaaagtttttcaaggaaatcaagctcctt 660
|||||
Db 4831 GTGGACTCTCTTTTTCAGTAATGCAAAAGTAAAGGTTTTCAGGGAATCAAGACTTCCTT 4890
QY 661 caaccctgtgtgaactctctagaaccacgcttactgactcgcttactccttcaatcaacc 720
|||||
Db 4891 CACACCTGTGTGAACCTCTTAAGACCCAGCTTACTGACTCGCTACCTTGAATTCACCC 4950
QY 721 ccgaagttggtgacacacattgcccctgaagatgagatgcttggctgacagacagga 780
|||||
Db 4951 CCAGAGTTGGGTGCACACAGATTCGCCCTGAGAGATGAGAGTTCTGGCTGCAGGACACAGA 5010
QY 781 cctctactgactcgagcc 798
|||||
Db 5011 CCTTACTGAGGGTGCC 5028

RESULT 8
AR112722
LOCUS AR112722 5035 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6130203.
ACCESSION AR112722
VERSION AR112722.1 GI:14092622
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Voorberg,J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 6130203-A 1 10-OCT-2000;
FEATURES
source 1..5035
location/Qualifiers
BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2.7e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaaggtgagatctgttggcaccatgatttcaagcagcaagaccaggtgccc 60
|||||
Db 4231 GATCAAGGTGAGATCTGTGGCACCAATGATTTTCACGGCATCAAGACCAGGTGCCCG 4290
QY 61 tcagaagttctccagcctctacatctctcagtttcaatcagatgaatcttgaatgaa 120
|||||
Db 4291 TCAGAAGTTCTCCAGCCTTACATCTCTCAAGTTTATCATCATGATATGCTTGATGGGAA 4350
QY 121 gaagttgcagacttaccgaagaattccactggaacttaagtctctcttggcaatg 180
|||||
Db 4351 GAAGTGGCAGACTTATTCAGGAATTCACCTGGAACCTTAATGCTCTTGGGAATGT 4410
QY 181 gattcatctgggataaacaacaatattttaaccctccaattatgtcgtatataccg 240
|||||
Db 4411 GATTTCATCTGGGATAAACAACAATATTTTAAACCTCCAAATATATGCTCGATATATCCG 4470
QY 241 ttggaccaccaactcatatagatctcgagacactcttgcagatgagatgaggtgtga 300
|||||
Db 4471 TTGGACCACTCATATATAGCATTTGCAAGCATCTTGGCATGAGGTGATGGGTGTGA 4530
QY 301 tttaaatagttgcagacatgcaatgggaatggaatgaagaacaatatacagatgcaagat 360
|||||
Db 4531 TTTAATAGTTGCAGCATGCCATTTGGGAATGAGATGAAGCAATATACATGACACAT 4590
QY 361 tactgttcatctacttaaccaataatgttgcacactggtctcctcaaaagctcgact 420
|||||

Db 4591 TACTGTTTACCTACTTACCAATATGTTTGCACACCTGCTCTCTCCAAAAGCTCGACT 4650
QY 421 taactccaagggagagatgaatgctgagacctcaggtgaatgaatcaaaagatggtc 480
|||||
Db 4651 TCACCTCCAAAGGAGAGATTAATGCTGAGACCTCAGCTGATATATCCAAAAGAGTGACT 4710
QY 481 gcaagtgcactccgaagaacaaatgaagtaacagagatgaactcagaggaatgaatc 540
|||||
Db 4711 GCAAGTGAACCTCCGAAGACATTAAGTCAAGAGATTAATGCTGAGGATGAATTC 4770
QY 541 tctgttaccagcagatgataatgaagagttcctcaatcctccagcagtcagaatgacata 600
|||||
Db 4771 TCTGCTTACCAAGCATATATGTAAGAGATTCCTCATCTCCAGCAGTCAGATGSCATCA 4830
QY 601 gtgagctctcttttttcaagaatgcaagtaaaagtttttcaaggaaatcaagctcctt 660
|||||
Db 4831 GTGGACTCTCTTTTTCAGTAATGCAAAAGTAAAGGTTTTCAGGGAATCAAGACTTCCTT 4890
QY 661 caaccctgtgtgaactctctagaaccacgcttactgactcgcttactccttcaatcaacc 720
|||||
Db 4891 CACACCTGTGTGAACCTCTTAAGACCCAGCTTACTGACTCGCTACCTTGAATTCACCC 4950
QY 721 ccgaagttggtgacacacattgcccctgaagatgagatgcttggctgacagacagga 780
|||||
Db 4951 CCAGAGTTGGGTGCACACAGATTCGCCCTGAGAGATGAGAGTTCTGGCTGCAGGACACAGA 5010
QY 781 cctctactgactcgagcc 798
|||||
Db 5011 CCTTACTGAGGGTGCC 5028

RESULT 9
LOCUS I77105 6999 bp DNA PAT 03-APR-1998
DEFINITION Sequence 1 from patent US 5693499.
ACCESSION I77105
VERSION I77105.1 GI:3013259
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6999)
AUTHORS Yonemura,H., Tajima,Y., Sugawara,K. and Masuda,K.
TITLE Process for preparing human coagulation factor VIII protein complex
JOURNAL Patent: US 5693499-A 1 02-DEC-1997;
FEATURES
source 1..6999
location/Qualifiers
BASE COUNT 2224 a 1503 c 1480 g 1792 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 6999;
Best Local Similarity 100.0%; Pred. No. 2.9e-234;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gataaagtgagatctgttggcaccatgatttcaagcagcaagaccaggtgccc 60
|||||
Db 6210 GATTAAGGTGAGATCTGTGGCACCAATGATTTTCACGGCATCAAGACCAGGTGCCCG 6269
QY 61 tcagaagttctccagcctctacatctctcagtttcaatcagatgaatcttgaatgaa 120
|||||
Db 6270 TCAGAAGTTCTCCAGCCTTACATCTCTCAAGTTTATCATCATGATATGCTTGATGGAA 6329
QY 121 gaagttgcagacttaccgaagaattccactggaacttaagtctcttggcaatg 180
|||||
Db 6330 GAAGTGGCAGACTTATTCAGGAATTCACCTGGAACCTTAATGCTCTTGGCAATG 6389
QY 181 gattcatctgggataaacaacaatattttaaccctccaattatgtcgtatataccg 240
|||||
Db 6390 GATTTCATCTGGGATAAACAACAATATTTTAAACCTCCAAATATATGCTCGATATATCCG 6449
QY 241 ttggaccaccaactcatatagatctcgagacactcttgcagatgagatgaggtgtga 300
|||||

|||||
Db 6450 TTGGACCACTATTATAGACTTCGACACTCTTCGATGGAGTTGATGGGCTGTGA 6509
Qy 301 tttaaatagttgacgatgcatggtggaatgaaagaaatatacaaatgacacagat 360
Db 6510 TTTAAATAGTTGACGACATGCCATTGGGAAATGAGAGTAAGACATATCAGATGACAGAT 6569
Qy 361 tactgcttcaacttaacaaatagtgttgcacactggtctctcttaaaagctgact 420
Db 6570 TACGCTTCATCCTACTTACCAATATGTTGCCACCTGGCTCTCTCAAAAGCTCGACT 6629
Qy 421 tcacctccaaggaggagaaatgctgtagaacctcaagtaataatacaaaaggtgact 480
Db 6630 TCACCTCCAAAGGAGGATATGCTGGAGACTTCAGGTATATCCAAAAGAGTGGCT 6669
Qy 481 gcaagtgaactccaagaagacatgaatgacaggaagtaactaactcaaggagtaaatc 540
Db 6690 GCAAGTGAATCCAGAAAGAAATGAAGTACAGAGTACTCTCAGGAGTAATAATC 6749
Qy 541 tctgcttaccagcatgtatgtgaaggagttccctcatctccagcagtaagaatggcatca 600
Db 6750 TCTGCTTACCGACATGATGTGAAGAGATTCCTCATCTCCAGCATCAAGATGGCATCA 6809
Qy 601 gtgagctctcttttttcaagaatgcaagaatgaaagttttcagggaatacaagctcctt 660
Db 6810 GTGGACTCTCTTTTTCAGAAATGCAAAAGTAAAGTTTTCAGGAAATCAAGACTCCTT 6869
Qy 661 cacacctgtgtgaactctctagaaccacgcttactgaactcgtactcctgaattcaacc 720
Db 6870 CACACCTGTGGGAGACTCTCTAGACCCACGTTACTGACTGCTTCAATTCACACC 6929
Qy 721 ccagaagttgggtgacacagatggtccctgaagatgaggtttctggtcgtgagagcaagga 780
Db 6930 CCAGAGTTGGGTGACACAGATTGCCCTGAGATGAGGTTCTGGCTGGCAGGACAGGA 6989
Qy 781 cctctactga 790
Db 6990 CCTCTACTGA 6999

RESULT 10
LOCUS 127063 7056 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 1 from patent US 5563045.
ACCESSION I27063
VERSION I27063.1 GI:1817839
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7056)
AUTHORS Pitman,D., Rehemtulla,A., Wozney,J.M. and Kaufman,R.J.
TITLE Chimeric procoagulant proteins
JOURNAL Patent: US 5563045-A 1 08-OCT-1996;
FEATURES
source 1..7056
BASE COUNT 2232 a 1517 c 1491 g 1816 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 7056;
Best Local Similarity 100.0%; Pred. No. 2.9e-234;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 gatcaagtgatgactgtgacccaatgattatcaagcgcataaagaccagggtgccc 60
Db 6267 GATCAAGTGATGACTGTGGACCAATGATATTTCACGGCATCAAGACCCAGGGTCCCG 6326
Qy 61 tcgaagttctcagccttacaatctcgaattatcatcatgataagcttgatggaa 120
Db 6327 TCAGAAGTTCTCAGCCTTACATCTCTCACTTATCATCATGATGATGATGATGAGGA 6386

Qy 121 gaagtgcagacttcatcggaatccactggaacttaatgctctcttggcaatgt 180
Db 6387 GAGTGGCAGACTTATTCGAGAAATTCACACTGGAACTTAATGCTTCTTGGCAATGT 6446
Qy 181 ggaatcatctggagataaacaacaatattttaacctcaatlatgtcgtacatcg 240
Db 6447 GGATTCATCTGGAGATAAACACAAATATTTTAACTCCCAATATATGCTGATACATCG 6506
Qy 241 ttggaccccaactcattatagacttgcagacactcttcgaagttgatgagcttga 300
Db 6507 TTTGGACCCCAACTATTATAGCATTTGACATTCGACACTCTTCGATGAGATTATGGGCTGTA 6566
Qy 301 tttaaatagttgacgatgcatggtggaatgagagtaaaagcaatacagatgacagat 360
Db 6567 TTTAAATAGTTGACGACATGCCATTGGGAATGAGAGTAAAGCAATATCAGATGACACAGAT 6626
Qy 361 tactgcttcaacttaacaaatagtgttgcacactggtctctcttaaaagctgact 420
Db 6627 TACTGCTTCATCCTACTTACCAATATGTTTGGCACCCTGGTCTCTTCAAAAGCTCGACT 6686
Qy 421 tcacctccaaggaggagaaatgctgtagaacctcaagtaataatacaaaaggtgact 480
Db 6687 TCACCTCCAAAGGAGGAGTATGCTGGAGACTTCAGGTATATCCAAAAGAGTGGCT 6746
Qy 481 gcaagtgaactccaagaagacatgaatgacaggaagtaactaactcagggagtaaatc 540
Db 6747 GCAAGTGAATCCAGAAAGAAATGAAGTACAGAGTACTCTCAGGAGTAATAATC 6806
Qy 541 tctgcttaccagcatgtatgtgaaggagttccctcatctccagcagtaagaatggcatca 600
Db 6807 TCTGCTTACCGACATGATGTGAAGAGTTCCTCATCTCCAGAGTCAAAAGATGCCATCA 6866
Qy 601 gtgagctctcttttttcaagaatgcaagaatgaaagttttcagggaatacaagctcctt 660
Db 6867 GTGGACTCTCTTTTTCAGAAATGCAAAAGTAAAGTTTTCAGGAAATCAAGACTCCTT 6926
Qy 661 cacacctgtgtgaactctctagaaccacgcttactgaactcgtactcctgaattcaacc 720
Db 6927 CACACCTGTGGGAGACTCTCTAGACCCACGTTACTGACTGCTTCAATTCACACC 6986
Qy 721 ccagaagttgggtgacacagatggtccctgaagatgaggtttctggtcgtgagagcaagga 780
Db 6987 CCAGAGTTGGGTGACACAGATTGCCCTGAGATGAGGTTCTGGCTGGCAGGACAGGA 7046
Qy 781 cctctactga 790
Db 7047 CCTCTACTGA 7056

RESULT 11
LOCUS 105404 7272 bp PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent EP 0265778.
ACCESSION 105404
VERSION 105404.1 GI:591046
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7272)
AUTHORS Sarver,N. and Prohan,W.
TITLE Factor VIII-C analogs
JOURNAL Patent: EP 0265778-A1 3 04-MAY-1988;
FEATURES
source 1..7272
BASE COUNT 2271 a 1591 c 1533 g 1877 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 7272;
Best Local Similarity 99.4%; Pred. No. 3e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	gataaagtgatctgttgcacaaatgattatccagcatcaagaccagggtgccg	60
Db	6267	GATCAAGGTGATCTGTGGCACCAGATGATTATTCAGGATCAAGACCAGGCTGCCG	6326
QY	61	tcagaagttccagcctcacatctctcaagttatcatcatgatagtcttgatggaa	120
Db	6327	TCAGAAATTCACAGCTCTACATCTCTCAGTTATCATCATGTATAGTCTTGATGGAA	6386
QY	121	gaagtcgaactatccaggaatccacatggaacttaagtctctcttggaagt	180
Db	6387	GAAGTCGACACTTATTCAGAGAAATTCACAGGACCTTATAGTCTCTTGACATGT	6446
QY	181	ggaatcctcgtggaataaacaacataatttaacccccaattatgctcgatataccg	240
Db	6447	GGAATCATCTGGGATTAACACATATTTTAACTCCATATTTGCTCGATACATCCG	6506
QY	241	tttgacccaactcatatagcaatctgcagcactcttcgcatgagtgatgagctgta	300
Db	6507	TTTGCACCCCACTCATATATATGATTCGACACTCTTCGATGAGTGTGATGGGCTGTGA	6566
QY	301	tttaaatagttgagagatgcatggaatggaatggaatgaagcaatatacgaatgcaagat	360
Db	6567	TTTAAATAGTTGAGCATGCTGGAATGAGAGATTAAGCAATATCAGATGCACAGAT	6626
QY	361	tactgtcactacacttataccaataatgttgcacactgctctcctccaagctgact	420
Db	6627	TACTGCTTACTACTTATACCAATATGTTTGACACTGCTCTCTCTTCAAAAAGCTCGACT	6686
QY	421	tcacctcccaaggaggaatgagctgagaccccaagtgaataatccaaagagtgact	480
Db	6687	TCACCTCCCAAGGAGAGATGCTCGAGACTCTCAGTGGAATATCCAAAAGAGTGCT	6746
QY	481	gcaagtggaactccagaagacaatgaagtcacaggaatgaactactcagggagtaaatc	540
Db	6747	GCAAGTGGACTTCAGAAACATGAATGCACAGAGATTAATCTCAGAGGAGTAATAATC	6806
QY	541	tctgttaccagcatgatagtgaagagttcctcactcctcagcaagtcagaatgagcatca	600
Db	6807	TCGTCTTACACAGATGTATGTGAAGAGTTCCTCATCTCCACAGATCAAGATGGCCATCA	6866
QY	601	gtggagctctcttttccagaatggaagtaagtttccagggaaatcaagactcctt	660
Db	6867	GTGGAGCTCTCTTTTCCAGAAATGCAAAAGTAAAGGTTTTCAGGAAATTCAGACTCTCT	6926
QY	661	caacactgtgtgaactcctcctagaccacacgcttactgactgctcactcgaattcacc	720
Db	6927	CACACTGTGTGAACCTCTAGACCCACACGCTTACTGACTGCTTCCATTCACATTCACCC	6986
QY	721	ccaagtgagtgagcaacagattgcccgaaggaatgaggttctgagctgagagcagaga	780
Db	6987	CCAGAGTTGGTGACACAGATGCTGAGATGAGAGGTCTGGGCTCGAGGACAGAGA	7046
QY	781	cctctactgactcagagcc	798
Db	7047	CCTCTACTGAGGTGGCC	7064

RESULT 12
 LOCUS AX052730 8967 bp DNA
 DEFINITION Sequence 1 from Patent WO0071714.
 ACCESSION AX052730
 VERSION AX052730.1 GI:12226920
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 8967)
 AUTHORS Saenko,E.L. and Strickland,D.K.
 TITLE Methods of reducing factor viii clearance and compositions therefor
 JOURNAL Patent: WO 0071714-A 1 30-NOV-2000;

		FEATURES		The American National Red Cross (US)	
		SOURCE		Location/Qualifiers	
		sig-peptide		1..8967	
		CDS		/db_xref="taxon:9606"	
				110..166	
				110..7165	
				/note="unnamed protein product"	
				/codon_start=1	
				/protein_id="CAC21811.1"	
				/db_xref="GI:12226921"	
				/translation="MQIEITSCFFLCILRFCSATRRYYLGAVELSMWVQSDGELP	
				VDARPPRRVSPSPSTSVYKKTLTPEFDHLFNTAKPPPMGLMTOAEVDT	
				VITLTKMSPHSYLHAGVSYKMSKASGADYDQTSOREKEDKVRGSGHTYVWVL	
				KENPMASDPLCLTSLYSLVVDYKDLNSGLTALCYCRSGSLAKETQTLHLFTLF	
				AVPEDEKSMHSEIKNSLMQORDASAAWAKMHTVMGVNRSPLGLGHRKSVYWH	
				IGMGTPEVHSIFLEHGTFLVRNRHQAISLSPITFLAOTLMDLGOFLFCHISSH	
				QHDMEAVYVVDSCPEEPOLRMKNNEEADYDLDLQSEMDVVRFPDDNSPFIORS	
				VAKHPRTWVHYIAAEEEDMDYAPLVAPDDRYSKSOYLNNQFORIGRKYKVRPMAY	
				TDEFFKRELAIOHRSGLTGLYGEVDLTLLIFKNQASRPYNIYPHGTDVRLYSR	
				RLKGVNHLKDFPLRPEITRYKWTYVEDGPKSDPKLTRYISFVNMEDLASGL	
				IGPLLYIKESVDQRNQIMSDKRNVLFSVPEDNSWYLTENIQRFLNPAGVOLD	
				PEFQASNMHSINGYVDSLSQCLHEVAVYILSIGAOTFLSYFEGYFEKHMV	
				YEDTLTLPESGETVEKSMENPGMLTICGNSQFRRNGMTALTKVSCDKNGDYED	
				YEDISAYLTKNNALTEPSEFSQNSRHSRROKQFNATTPENDIEKTDPMFAHPTM	
				PKIOWNSSDLMLLROSPPHGLSLSDQEAKEYFPSPDRSGAIDSNNSLSBMTF	
				RPLQHRHGDVFTTPESGLRLNEKGLTATLKLKLDKRVSTSNLLISTIPSDMA	
				AGTDNITSLSLPPSPVHYDSQDPTLFGKSSPLTBSGPLSLSENNDSKLLSEGLM	
				NSQESSWKNVSTESGRLEFKGRAGPALTKDNALFVYSISLTKNTSNMSATNR	
				KTHIDGSLILENSPLYWONIIEBDEFEKVPPLIDRMILMKNATALRLNMSNTY	
				SKNMEWVOOKKEGPIEPDQNPDMSEFKMLFLPESARNTORTHGKNSLNSGQSPK	
				OLVSLGPEKEVEQONFLSEKNKYVVGKGFETQVGLKEMVFPSSRLFLPINDLHEN	
				NTHNOEKIOEIEKEKTELLQENVYLPQHTYGTGNPKNLPLSTRONVBSIDGA	
				YAVLQDFRLANDSTNRKTHAHSEKSGEENLELQNOTQIYEKACTRIISPT	
				SOONFVORSKRLKOFRLPELETELEKRIYDVTSTQMSKNNKHLPTSLQIDNE	
				KEKGATIOSPLDSCLTRSHSIPQANSPLPIAKVSSPISIRIYLTVRFQONSHP	
				AASYRRKDSVQESHFLQAGAKNNLSLILTEMDGQREVSIGTSATNSVTKV	
				ENIVLPDLPKTSKGVLLPKVHIYOKDLPETNGSGPLHDLVEGSLTGTGAI	
				KMEANRPGKVPFLRVATSSAKTPSKLIDPLAMDHVGTQIPKEMSOEKSPEKTA	
				FRKKDTILSNACESHNAIINEQNKREIETNAKQRTERLCSQNPVLRKQRE	
				ITRTTLOSDEIDDDTTSVEKKKEDFPIYEDENQSRSSQKTRHFLAABRLM	
				DYGMSSPHYLNRRAQSGVPOFKKVPQKVVLEDSSTQPLLYGELNHEHGLGILYRA	
				EVEDNIMVTFRNQAQSPYSFSLISYEEDQGLCAPRNFKPVKNSTKTYFAKVOHM	
				APRKDFEDFCAMAYFSDVDLEKDHSGLLPLVCTTNLANHAGSTQVVOEFLPT	
				IPDETSTWYETEMNERCAPQNIOMEDPFEENRPHAINVYINDTIDPLGAMADOR	
				IRKYLISMSNENIHSIHRSGHVFTYRKBEKMLNLYIPVFEVEMLPKACIMR	
				VECLIEHHLHAGSTLFLVYSNCCQPLDMASGHLRDFQITASGYQWAPLRLHY	
				SGSINAMSTRKEPSPWIKVDLAPMIHGIKQGAAROKFESSLYISQFIIVSIDGRKW	
				TYRGNSTGLMVEFGWVDSGIKHNIHPPIIARYIRLPHYISIRSTLRMLGCDL	
				NCSMPLGMSKASIDAOITASSYFMNFAATPSKARLHLGRRSAMPPOVNNPKEM	
				LOYDFQTKMKVGTGTVGKSLTSLTSYVEFLISSQDSQHWOTLFPQNKVAVPQNG	
				DSFTPVNSLDPLRLRYLRHHPQSVNHDIALRMEVLGCEADLYT	

mat_peptide 2841 a 1898 c 1833 g 2395 t
 BASE COUNT 2841 a 1898 c 1833 g 2395 t
 ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 8967;
 Best Local Similarity 99.4%; Pred. No. 3, 1e-234;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	gataaagtgatctgttgcacaaatgattatccagcatcaagaccagggtgccg	60
Db	6376	GATCAAGGTGATCTGTGGCACCAGATGATTATTCAGGATCAAGACCAGGCTGCCG	6435
QY	61	tcagaagttccagcctcacatctctcaagttatcatcatgatagtcttgatggaa	120
Db	6436	TCAGAAATTCACAGCTCTACATCTCTCAGTTATCATCATGTATAGTCTTGATGGAA	6495
QY	121	gaagtcgaactatccaggaatccacatggaacttaagtctctcttggaagt	180
Db	6496	GAAGTCGACACTTATTCAGAGAAATTCACAGGACCTTATAGTCTCTTGACATGT	6555

QY 181 ggaatcctcgtggaataaacaacaaatatttaacccctccaaattatgtctgatacatccg 240
 DB 6556 GGATTCATCTGGGATTAACAAACAAATATTTTAAACCTCCAAATATTTCTGATACATCCG 6615
 QY 241 ttgaccccaactcattatagcattcgcagacactcgcagatgagttgagctgtga 300
 DB 6616 TTGGACCCCACTCATATTAATGATTCGACACTCTTCGCAATGAGTTGATGGCTGTGA 6675
 QY 301 tttaaatagttgacagatcgcattgaggaatgagaaagcaatatacagatgacagat 360
 DB 6676 TTAAATAGTTGACAGATGCGCATTTGGGAATGAGAGATAAAGCAATATCAGATGACAGAT 6735
 QY 361 tactgtcactcacttaccataatgtttgcacactgtctctccaaagctgact 420
 DB 6736 TACTGCTTCATCTCACTTACCAATATGTTTGGCCACTGCTCTCTTCAAAAGCTGACT 6795
 QY 421 tcaactcagaaggaggaatgatacctcgcagacccagtgatataatccaaaggtgact 480
 DB 6796 TCACCTCCAAAGGAGAGATATGCTGAGACCTCAGTGAATATCCAAAGAGTGGCT 6855
 QY 481 gcaagtgactccagaagaacatgaaagtcacagagagtaactcagaggagtaaatc 540
 DB 6856 GCAAGTGACTCTCCAGAGACAAATGAAGTCAAGAGTACTACTCAGGAGTAAATC 6915
 QY 541 tctgttaccagcatgtatgtgaaaggagttcctatctccagacagtaagaatgcatca 600
 DB 6916 TCTGCTTACCAAGCATGATATGTAAGAGAGTCTCTCATCTCCAGACGTCAGATGGCATCA 6975
 QY 601 gtgagctccttttccagaatgcaagaatgagtttttcaaggaaatcagactcctt 660
 DB 6976 GTGAGACTCTCTTTTTCAGAAATGCAAGTAAGTTCAGGGAATTCAGACTCCTT 7035
 QY 661 caacactgtgtgaactctctagaccacagcttactgactcgtactcgtgaatcacc 720
 DB 7036 CACACCTGTGTGAACCTCTAGACCCACCGCTTACTGACTGCTTCAATTCACCC 7095
 QY 721 ccaagattggtgacacagattgacctgaagatgaggttctgagctgcagagcaaga 780
 DB 7096 CCAGAGTTGGGTGCACCAAGATTGCCCTGAGAGTGGAGTTCTGGGCTCGAGGCACAGGA 7155
 QY 781 cctctactgactgagcc 798
 DB 7156 CCTTACTGAGGTTGGCC 7173
 RESULT 13
 LOCUS 171409 8967 bp DNA PAT 03-APR-1998
 DEFINITION Sequence 1 from patent US 5681746.
 ACCESSION 171409
 VERSION 171409.1 GI:3007544
 KEYWORDS
 ORGANISM Unknown.
 SOURCE Unknown.
 REFERENCE 1 (bases 1 to 8967)
 AUTHORS Bodner, M., De Polo, N.J., Chang, S., Hsu, D. Chi-Tang and Respass, J.G.
 TITLE Retroviral delivery of full length factor VIII
 JOURNAL Patent: US 5681746-A 1 28-OCT-1997;
 FEATURES
 source 1..8967
 /organism="Unknown"
 BASE COUNT 2841 a 1898 c 1833 g 2395 t
 ORIGIN
 Query Match 79.0%; Score 790; DB 6; Length 8967;
 Best Local Similarity 99.4%; Pred. No. 3,1e-234;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 gatcaagtgatcgtgtgtgacccaatgatatccagcgcacccaagaagcaggtgccc 60
 DB 6376 GATCAAGTGATGATCTGTGGACCAATGATATTCACGCGCATCAAGACCCAGGTCGCG 6435

QY 61 tcagaagttctcagcctctacatctctcaagttatcatcatgataagttgagga 120
 DB 6436 TCAGAAGTTCTCAGCCTCTACATCTCTCAAGTTATCAATCATATGATCTGATGGGA 6495
 QY 121 gaagtgcagactatcaggaagaattccactggaacttaagttcttcttggcaatgt 180
 DB 6496 GAAGTGGCACTTATCGAGGAATTCACACTGGAACTTAATGCTCTCTTGGCAATGT 6555
 QY 181 ggaatcctcgtggaataaacaacaaatatttlaacccctccaaattatgtctgatacatccg 240
 DB 6556 GGATTCATCTGGGATTAACAAACAAATATTTTAAACCTCCAAATATTTCTGATACATCCG 6615
 QY 241 ttgaccccaactcattatagcattcgcagacactcgcagatgagttgagctgtga 300
 DB 6616 TTGGACCCCACTCATATTAATGATTCGACACTCTTCGCAATGAGTTGATGGCTGTGA 6675
 QY 301 tttaaatagttgacagatcgcattgaggaatgagaaagcaatatacagatgacagat 360
 DB 6676 TTAAATAGTTGACAGATGCGCATTTGGGAATGAGAGATAAAGCAATATCAGATGACAGAT 6735
 QY 361 tactgtcactcacttaccataatgtttgcacactgtctctccaaagctgact 420
 DB 6736 TACTGCTTCATCTCACTTACCAATATGTTTGGCCACTGCTCTCTTCAAAAGCTGACT 6795
 QY 421 tcaactcagaaggaggaatgatacctcgcagacccagtgatataatccaaaggtgact 480
 DB 6796 TCACCTCCAAAGGAGAGATATGCTGAGACCTCAGTGAATATCCAAAGAGTGGCT 6855
 QY 481 gcaagtgactccagaagaacatgaaagtcacagagagtaactcagaggagtaaatc 540
 DB 6856 GCAAGTGACTCTCCAGAGACAAATGAAGTCAAGAGTACTACTCAGGAGTAAATC 6915
 QY 541 tctgttaccagcatgtatgtgaaaggagttcctatctccagacagtaagaatgcatca 600
 DB 6916 TCTGCTTACCAAGCATGATATGTAAGAGAGTCTCTCATCTCCAGACGTCAGATGGCATCA 6975
 QY 601 gtgagctccttttccagaatgcaagaatgagtttttcaaggaaatcagactcctt 660
 DB 6976 GTGAGACTCTCTTTTTCAGAAATGCAAGTAAGTTCAGGGAATTCAGACTCCTT 7035
 QY 661 caacactgtgtgaactctctagaccacagcttactgactcgtactcgtgaatcacc 720
 DB 7036 CACACCTGTGTGAACCTCTAGACCCACCGCTTACTGACTGCTTCAATTCACCC 7095
 QY 721 ccaagattggtgacacagattgacctgaagatgaggttctgagctgcagagcaaga 780
 DB 7096 CCAGAGTTGGGTGCACCAAGATTGCCCTGAGAGTGGAGTTCTGGGCTCGAGGCACAGGA 7155
 QY 781 cctctactgactgagcc 798
 DB 7156 CCTTACTGAGGTTGGCC 7173
 RESULT 14
 LOCUS HSFVIIIR 8967 bp mRNA PRI 21-MAR-1995
 DEFINITION Human mRNA for factor VIII.
 ACCESSION X01179
 VERSION X01179.1 GI:31498
 KEYWORDS factor VIII; signal peptide.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 8967)
 AUTHORS Wood, W.I., Capon, D.J., Simonsen, C.C., Eaton, D.L., Gitschier, J.,
 Keyt, B., Seeburg, P.H., Smith, D.H., Hollingshead, P., Wion, R.L.,
 Delwart, E., Tuddenham, E.G.D., Vohar, G.A. and Lawn, R.M.
 TITLE Expression of active human factor VIII from recombinant DNA clones
 JOURNAL Nature 312 (5992), 330-337 (1984)
 MEDLINE 85061548
 COMMENT Data kindly reviewed (20-MAR-1986) by W. Wood.

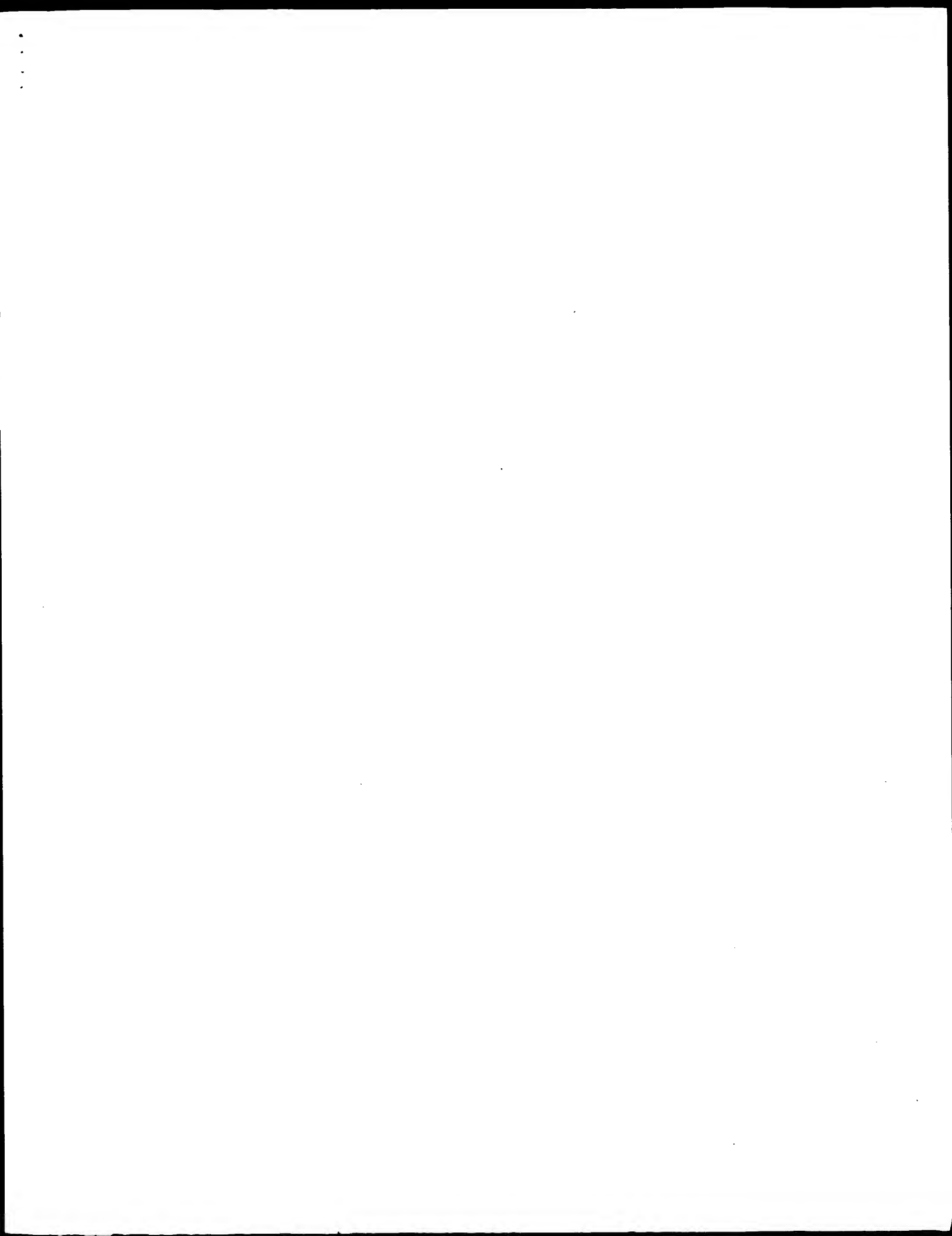
ORIGIN

Query Match 79.0%: Score 790; DB 6; Length 9009;

Best Local Similarity 99.4%: Pred. No. 3,1e-234; Mismatches 793; Conservative 0; Indels 5; Gaps 0;

```
QY 1 gataaagtgatctgtgtgacccaatgattatcacgacatcaagaccaggtgccg 60
Db 6417 GATCAAGGTGATCTGTGGSCACCAATGATATTCACGGCATCAAGACCAGSGTGCCG 6476
QY 61 tcagaagttccagcctctacatctcagttatcatatglatagtccttgaggaa 120
Db 6477 TCAGAAGTTCTCCAGCCCTCAGATCTCAGTTATCATCATATAGTCTTGATGGAA 6536
QY 121 gaagtggacactatcgaggaatctcactggacctaagtctctttggcaatgt 180
Db 6537 GAAGTGGACACTTATCGAGGAATTCACCTGAACCTTAATGGCTCTTTGGCAATGT 6596
QY 181 ggaatcactggataaacaacaatattttaaccctccaattatgtctgataatc 240
Db 6597 GGATTTCATCTGGATTAACAACATATTTTAACCTCCAAATTTGCTGATACATCCG 6656
QY 241 ttgcaccccaatcatatagcaltcgagcactcttcgcatggagttga tggcttga 300
Db 6657 TTTGCACCCCACTCATATATAGCATTCGACACACTCTTCGATGAGTTGATGGCTGTGA 6716
QY 301 ttttaaatgtgcaagcattgcattgggaatggagatgaagaacaatcagatgcagat 360
Db 6717 TTTAAATAGTTGACAGCATGCCATTGGGAATGGAGTAAGCAATATACAGATGCACGAT 6776
QY 361 taagcttcaatcacttaccatatagtttgacacctgtgtccttcaaaagctgact 420
Db 6777 TACTGCTTCACTCTACTTTTACCAATATGTTGGCACCTGGTCTCTTCAAAAGCTGACT 6836
QY 421 tcaacctcaaggaggaglaatgcttggagacactcaggtgaataatccaaaagttgct 480
Db 6837 TCACCTCCAAAGGAGGAGTATGCTGAGACCTCAGGTGAATATCCAAAGAGTGCT 6896
QY 481 gcaagtgtgacttccagaagacaatgaagtcaagaagttaactactcagggagtaaatc 540
Db 6897 GCAAGTGGACTTCAGAGAACAAATGAAGTACACAGAGTAACACTCAGGAGTAAATC 6956
QY 541 tctgcttaccagcaatgatttgaagagttccatctccagcagttcaagatggcaatca 600
Db 6957 TCTCTTACCAACATGATATGAGAGGTTCTCTATCTCCAGCATGAAGATGGCCATCA 7016
QY 601 gtgactctcttttcaagaatggcaaglaaagttttcagggaaatcaagactcctt 660
Db 7017 GTGACTCTCTTTTTCAGAAATGCAAGTAAGGTTTTCAGGGAATCAAGACTCTCTT 7076
QY 661 cacaactgtgtgaactctctagaaccacgcttaactgaactcgtactccttgaatcaacc 720
Db 7077 CACACTGTGTGTAACCTCTAGACCCACCGTTACTGACTGCTTGAATTCACCC 7136
QY 721 ccagaagttgggtgacacagattgcccctgagatggaggttctgtgctgagagcacagga 780
Db 7137 CCAGAGTTGGGTGACCAATGCTCGAGAGATGAGAGTTCTGGCTGCGAGGACACAGGA 7196
QY 781 cctctactgaactgagcc 798
Db 7197 CCTCTACTGAGGGTGCC 7214
```

Search completed: January 17, 2002, 16:26:53
Job time: 13934 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:34:09 ; Search time 343.96 Seconds

(without alignments)
2492.514 Million cell updates/sec

Title: US-09-740-211-14_COPY_4000_4999

Perfect score: 1000

Sequence: 1 gacaaagttgacgtgtgtg.....cgcagctgcctgcagacat 1000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: N_Geneseq_1101.*
2: /SID2/gcgdata/geneseq/geneseq/NA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq/NA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq/NA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq/NA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq/NA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq/NA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq/NA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq/NA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq/NA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq/NA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq/NA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq/NA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq/NA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq/NA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq/NA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq/NA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq/NA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq/NA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq/NA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq/NA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq/NA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	4999	21	AD000122
2	1000	100.0	4999	22	AD008613
3	954.4	95.4	11933	21	AD000121
4	954.4	95.4	11933	22	AD008612
5	870.8	87.1	7964	22	AA84647
6	790	79.0	4275	10	AA90654
7	790	79.0	4629	20	AA88293
8	790	79.0	4670	19	AAV23339
9	790	79.0	4832	19	AAV19581
10	790	79.0	4832	19	AAV15338
11	790	79.0	5035	18	AA69811

12	790	79.0	5094	21	AAA49231	DNA construct H5Q/
13	790	79.0	7056	15	AA066615	Sequence of human
14	790	79.0	8967	17	AA731031	Factor-VIII full-1
15	790	79.0	8967	22	AA87526	Human factor VIII
16	790	79.0	8975	6	AA50054	Human factor VIII
17	790	79.0	8975	21	AA238604	Human full-length
18	790	79.0	9009	14	AA050185	Human factor VIII
19	790	79.0	9009	18	AA61548	Human factor VIII
20	790	79.0	9009	19	AA25810	Factor VIII:C (Arg
21	790	79.0	9009	19	AAV12112	Human factor VIII
22	790	79.0	9009	20	AAV12112	Human factor VIII
23	790	79.0	9009	22	AA690508	Human factor VIII
24	790	79.0	9029	22	AA690508	Human factor VIII
25	790	79.0	9068	19	AAV15359	Human factor VIII
26	790	79.0	9080	19	AAV19580	Human factor VIII
27	790	79.0	9164	20	AA82259	Human factor III e
28	790	79.0	9354	18	AA73164	Beta-domain delete
29	790	79.0	11846	20	AA82261	CDNA encoding huma
30	790	79.0	12022	20	AA82260	Factor VIII protei
31	790	79.0	12445	21	AA49232	Factor VIII protei
32	788.4	78.8	4616	9	AA81545	Vector H5QReNeo fo
33	788.4	78.8	4830	9	AA81544	Human factor VIII-
34	788.4	78.8	7059	9	AA81543	Human factor VIII-
35	788.4	78.8	8241	9	AA81439	Human factor VIII-
36	788.4	78.8	8241	9	AA81096	Factor VIII cDNA i
37	788.4	78.8	9009	19	AAV18884	CDNA sequence enco
38	787	78.7	4272	9	AA80447	Homo sapiens facto
39	787	78.7	4275	9	AA80446	Modified factor VI
40	787	78.7	4545	9	AA80444	Modified factor VI
41	787	78.7	7053	18	AA751357	Factor VIII:C codi
42	785.2	78.5	7440	7	AA60689	Sequence encoding
43	782	78.2	4629	16	AA676016	B-domain deleted F
44	764.4	76.4	7440	6	AA50375	DNA sequence enco
45	741	74.1	4373	20	AA82258	Beta-domain delete

ALIGNMENTS

RESULT 1	
AD000122	AD000122 standard; DNA; 4999 BP.
XX	
AC	AD000122;
XX	
DT	31-JUL-2000 (first entry)
XX	
DE	Recombinant adeno associated vector construct, pVmd.lcf8deltaB.
XX	
KW	Recombinant Adeno Associated Vector; RAAY; pVmd.lcf8deltaB; promoter;
KW	human factor VIII; hVfIII; EF1alpha; human elongation factor-1alpha;
KW	human growth hormone; hGH; ITR; Inverted terminal repeat; haemophilina;
KW	gene therapy; ds.
XX	
OS	Adeno associated virus.
XX	
PN	WO200023116-A1.
XX	
PD	27-APR-2000.
XX	
PF	19-OCT-1999; 99WO-US24495.
XX	
PR	20-OCT-1998; 98US-0104994.
PR	24-MAR-1999; 99US-0125974.
PR	30-JUL-1999; 99US-0364862.
XX	
PA	(AVTG-) AVIGEN INC.
XX	
PI	Conto LB, Colosi PC;
XX	
DR	WPI; 2000-339536/29.
XX	
PT	New recombinant adenovirus-associated vector, useful for gene therapy

to treat hemophilia, comprises at least a portion of Factor VIII
 operably linked to control sequence -

Example 9; Fig 6; 92pp; English.

The present DNA sequence is a recombinant adenovirus-associated vector, (RAV) construct pV4.1cF8deltaB. This expression vector comprises the promoter, first intron (-573 to +985) of human elongation factor-1alpha (EF1alpha) gene, human Factor VIII coding sequence (hFVIII) and a polyadenylation signal from human growth hormone (hGH). This sequence is inserted between the RAV inverted terminal repeat (ITR) regions. The hFVIII coding region comprises the heavy chain gene segment with the first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains and 5 amino acids from the N-terminus of the B domain. The light chain segment comprises the C-terminal 85 amino acids of B domain and the A3, C1 and C2 domains. Both the heavy and light chain segments are cloned into the same plasmid separated by 42 nucleotides coding for 14 residues of the B domain, that is deleted. This plasmid is operably linked to control sequences, that directs the transcription and translation of the Factor VIII gene. The adeno-associated viral vectors are used for gene therapy to treat haemophilia. This method allows prolonged expression of therapeutic levels of Factor VIII in vivo. The RAV are used for gene therapy, because of their broad host range, safety profile and duration of expression in the infected hosts.

Sequence 4999 BP; 1375 A; 1151 C; 1155 G; 1318 T; 0 other;

Query Match 100.0%; Score 1000; DB 21; Length 4999;

Best Local Similarity 100.0%; Pred. No. 5.6e-302;
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcaagtgatgactctgtggcaccatgatttcaagcgcacacccagggtgccg 60
 DB 4000 gatcaagtgatgactctgtggcaccatgatttcaagcgcacacccagggtgccg 4059
 QY 61 tcgaagttctcagcctctacatctcagttatcatcatcatgactgctgatggaa 120
 DB 4060 tcgaagttctcagcctctacatctcagttatcatcatcatgactgctgatggaa 4119
 QY 121 gaagtgacagactatcagaagaaatccatcggaaacttaagtctctctggcaagt 180
 DB 4120 gaagtgacagactatcagaagaaatccatcggaaacttaagtctctctggcaagt 4179
 QY 181 ggaatcatctggagataaacaacaatattttaacctcaattatgtcctgataactcg 240
 DB 4180 ggaatcatctggagataaacaacaatattttaacctcaattatgtcctgataactcg 4239
 QY 241 ttggccccaactcattatagacatctgcagcactcttcgcatgagtgatgagctgtga 300
 DB 4240 ttggccccaactcattatagacatctgcagcactcttcgcatgagtgatgagctgtga 4299
 QY 301 tttaaatagtgacagatcgtcatttggagagtgagagtaaaagcaatalcagatgcacagat 360
 DB 4300 tttaaatagtgacagatcgtcatttggagagtgagagtaaaagcaatalcagatgcacagat 4359
 QY 361 tactgtctcatctacttaccacaatagtgttgcacactgtgtcctctcaaaaagctgcact 420
 DB 4360 tactgtctcatctacttaccacaatagtgttgcacactgtgtcctctcaaaaagctgcact 4419
 QY 421 tcacctcaagggagagtaatgctgagagactcagtgtaataatccaaaagtgagct 480
 DB 4420 tcacctcaagggagagtaatgctgagagactcagtgtaataatccaaaagtgagct 4479
 QY 481 gcaagtgacttcagaagaacaatgaaagtcacagagtaactactcaaggagtaaatc 540
 DB 4480 gcaagtgacttcagaagaacaatgaaagtcacagagtaactactcaaggagtaaatc 4539
 QY 541 tcgcttaccagatctgagtgaagaggttcctcatctccagagagtaaatgagcatalc 600
 DB 4540 tcgcttaccagatctgagtgaagaggttcctcatctccagagagtaaatgagcatalc 4599
 QY 601 gtgactctcttttcagatgacaaagtaaggttttcagagaaatcaagactcctt 660

DB 4600 gtgacactctcttttcagatgacaaagtaaggttttcagagaaatcaagactcctt 4659
 QY 661 cacactgtgtgtaactctctagaccacccgttaactgaactgctactcctcgaatttacc 720
 DB 4660 cacactgtgtgtaactctctagaccacccgttaactgaactgctactcctcgaatttacc 4719
 QY 721 ccagagttggtgacacagatgctccctgagatgagaggttctgtgctgcagacagga 780
 DB 4720 ccagagttggtgacacagatgctccctgagatgagaggttctgtgctgcagacagga 4779
 QY 781 cctctactgactcagcctataaaggaatatttattcattgcaatagtgtgtgtt 840
 DB 4780 cctctactgactcagcctataaaggaatatttattcattgcaatagtgtgtgtt 4839
 QY 841 ttgtgtgcgcgcaggaacccctatgtatgagtgagtgagcactcctctcgcgcgc 900
 DB 4840 ttgtgtgcgcgcaggaacccctatgtatgagtgagtgagcactcctctcgcgcgc 4899
 QY 901 gctgcctcactgagcgcgcgcagacaaagtcgccgacgcgcgcgttgcgcgcgcgc 960
 DB 4900 gctgcctcactgagcgcgcgcagacaaagtcgccgacgcgcgcgttgcgcgcgcgc 4959
 QY 961 cctcagtgagc 1000
 DB 4960 cctcagtgagc 4999

RESULT 2
 AAD08613
 ID AAD08613 standard; DNA; 4999 BP.
 XX
 AC AAD08613:
 XX
 DT 04-SEP-2001 (first entry)
 XX
 DE Human factor VIII expressing RAV vector pV4.1cF8-B partial sequence.
 KW Recombinant adeno-associated viral vector; RAV; factor VIII; FVIII;
 KW blood clotting disorder; gene therapy; haemophilia A; human;
 KW pV4.1cF8-B; ds.
 OS Chimeric - Adeno associated virus.
 OS Chimeric - Homo sapiens.
 OS Chimeric - Unidentified.
 XX
 PM W0200145510-A1.
 XX
 PD 28-JUN-2001.
 XX
 PF 21-DEC-2000; 2000WO-US34925.
 XX
 PR 22-DEC-1999; 99US-0470618.
 XX
 PA (AVIG-) AVIGEN INC.
 XX
 PI Couto LB, Colosi PC, Qian X;
 XX
 DR WPI: 2001-417955/44.
 XX
 PT Treating blood clotting disorder, especially hemophilia in mammals, by
 PT administering recombinant adeno-associated vectors which express blood
 PT coagulation factor VIII -
 XX
 PS Claim 19; Fig 6; 90pp; English.
 CC The present invention relates to a method for treating a subject
 CC suffering from a blood clotting disorder. The method comprises
 CC administering a recombinant adeno-associated virion (RAV) comprising
 CC a nucleotide sequence encoding the light chain of factor VIII and a
 CC second recombinant adeno-associated virion comprising a nucleotide
 CC sequence encoding the heavy chain of factor VIII. The RAV vector is
 CC useful in gene therapy for treating haemophilia A in mammals, in

CC particular humans. The rAAV vector provides high level and long term
 CC expression of biologically active clotting factor VIII in vivo.
 CC The present sequence is pYM4.lcF8-B vector without the plasmid,
 CC backbone. It comprises human elongation factor 1 (EF1) promoter,
 CC B-domain deleted human factor VIII coding sequence, a poly A signal
 CC and AAV inverted terminal repeats (ITRs) at the ends. The vector
 CC encodes both the light and heavy chains of human factor VIII
 CC separated by 14 amino acids of the B domain.

XX Sequence 4999 BP; 1375 A; 1151 C; 1155 G; 1318 T; 0 other;

Query Match 100.0%; Score 1000; DB 22; Length 4999;

Best Local Similarity 100.0%; Pred. No. 5.6e-302;

Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 gatcaagtgatctgttggccaatgtattatcagcgctcaagaccgggtgccc 60
DB 4000 gatcaagtgatctgttggccaatgtattatcagcgctcaagaccgggtgccc 4059
OY 61 tcagaagttccagcctcatctcagtttatacatgatacttcttgatgagaa 120
DB 4060 tcagaagttccagcctcatctcagtttatacatgatacttcttgatgagaa 4119
OY 121 gaagtgagcagactatcgaggaattccatggaaccttaatgtctcttggcaatgt 180
DB 4120 gaagtgagcagactatcgaggaattccatggaaccttaatgtctcttggcaatgt 4179
OY 181 gattcatctggataaaacacatatttttaacctcaattatgtctgataatccg 240
DB 4180 gattcatctggataaaacacatatttttaacctcaattatgtctgataatccg 4239
OY 241 ttgcacccaactcatatagatctgcagcactcttcgcatgagtgatgagtgctgta 300
DB 4240 ttgcacccaactcatatagatctgcagcactcttcgcatgagtgatgagtgctgta 4299
OY 301 ttttaaatgttcagcagtcgacttgggaatgaggttaagaataatagatgacaaat 360
DB 4300 ttttaaatgttcagcagtcgacttgggaatgaggttaagaataatagatgacaaat 4359
OY 361 tactgttcatctactatttaccataatgtttgcacctgtctccttcaaaagctcgact 420
DB 4360 tactgttcatctactatttaccataatgtttgcacctgtctccttcaaaagctcgact 4419
OY 421 tcacctccaaggaggaatgatactgctgagacctcgaagtgaataatccaagaatgct 480
DB 4420 tcacctccaaggaggaatgatactgctgagacctcgaagtgaataatccaagaatgct 4479
OY 481 gcaagtgagacttcagaagaacatgaaatgcagaggaactactcagggagtaaatc 540
DB 4480 gcaagtgagacttcagaagaacatgaaatgcagaggaactactcagggagtaaatc 4539
OY 541 tctgcttaccagcatgtatgtgaaagagttcctcatctccaagtcgaagtgcacatca 600
DB 4540 tctgcttaccagcatgtatgtgaaagagttcctcatctccaagtcgaagtgcacatca 4599
OY 601 gtagactctcttttccaagaatgcaagtaaaagttttccaaggaatacgaactcctt 660
DB 4600 gtagactctcttttccaagaatgcaagtaaaagttttccaaggaatacgaactcctt 4659
OY 661 caaacctgtgtgaactctcttagaccacacgttactgactcgctactcgaattcacc 720
DB 4660 caaacctgtgtgaactctcttagaccacacgttactgactcgctactcgaattcacc 4719
OY 721 ccgaagttgggtgacacagatgcccctgaaatgagtggtctggtctcgagggcacagga 780
DB 4720 ccgaagttgggtgacacagatgcccctgaaatgagtggtctggtctcgagggcacagga 4779
OY 781 ccttactgactcgagcctataaagaaattattttatcttcatatagtggtgtgtt 840
DB 4780 ccttactgactcgagcctataaagaaattattttatcttcatatagtggtgtgtt 4839
OY 841 ttgtgtggtggccgacgaacccctagtgatgtgagtgccactcctctctgctgctc 900

```

```

DB 4840 ttgtgtggtggccgacgaacccctagtgatgtgagtgccactcctctctgctgctc 4899
OY 901 gctcgtcactgagggcgagccacaaagttgcgccagccggcttccggcgcg 960
DB 4900 gctcgtcactgagggcgagccacaaagttgcgccagccggcttccggcgcg 4959
OY 961 cctcaagtgaagcgagcgagcgagctgtcctgaggaact 1000
DB 4960 cctcaagtgaagcgagcgagcgagctgtcctgaggaact 4999

```

RESULT 3

AAD00121 standard; DNA; 11933 BP.

AAD00121; 31-JUL-2000 (first entry)

Recombinant adeno associated vector construct, PAAV-F8-1.

Recombinant Adeno Associated Vector; PAAV, PAAV-F8-1; human Factor VIII; hVYIII; hNF-3 albumin promoter; human elongation factor-1alpha; EF1alpha; human growth hormone; hGH; inverted terminal repeat; ITR; haemophilia; gene therapy; ds.

Adeno associated virus.

W0200023116-A1.

27-APR-2000.

19-OCT-1999; 99WO-US24495.

20-OCT-1998; 98US-0104994.

24-MAR-1999; 99US-0125974.

30-JUL-1999; 99US-0364862.

(AVIGEN-) AVIGEN INC.

Couto LB, Colosi PC;

WPI; 2000-339536/29.

New recombinant adenovirus-associated vector, useful for gene therapy to treat hemophilia, comprises at least a portion of Factor VIII operably linked to control sequence.

Example 2; Fig 5; 92pp; English.

The present DNA sequence is a recombinant adenovirus-associated vector, (rAAV) construct, PAAV-F8-1. This expression vector comprises the hNF-3 albumin promoter, the first intron (-573 to +985) of human elongation factor-1alpha (EF1alpha) gene, human factor VIII coding sequence (hVYIII) and polyadenylation signal from human growth hormone (hGH). This sequence is inserted between the AAV inverted terminal repeat (ITR) regions. The hVYIII coding region comprises the heavy chain gene segment with the first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains and 5 amino acids from the N-terminus of the B domain. The light chain segment comprises the C-terminal 85 amino acids of B domain and the A3, C1 and C2 domains. Both the heavy and light chain segments are cloned into the same plasmid separated by 42 nucleotides coding for 14 residues of the B domain, that is deleted. This plasmid is operably linked to control sequences, that directs the transcription and translation of the Factor VIII gene. The adeno-associated viral vectors are used for gene therapy to treat haemophilia. This method allows prolonged expression of therapeutic levels of Factor VIII in vivo. The rAAV are used for gene therapy, because of their broad host range, safety profile and duration of expression in the infected hosts.

Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;

Query Match 95.4%; Score 954.4; DB 21; Length 11933;
 Best Local Similarity 97.9%; Pred. No. 1.7e-287;
 Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

```

QY 1 gatcaagtgatgctgttgacccaatgattatcacgacatcaagaaccagggtgccg 60
   |||||||
DB 3975 gatcaagtgatgctgttgacccaatgattatcacgacatcaagaaccagggtgccg 4034
QY 61 tcaaaattccacgactctacatctcagttatcatatcatatgattatgattggaa 120
   |||||||
DB 4035 tcgaagttccacgactctacatctcagttatcatatcatatgattatgattggaa 4094
QY 121 gaagtgacagactatcgaggaattccactggaaacttaatgctctcttggaatgt 180
   |||||||
DB 4095 gaagtgacagactatcgaggaattccactggaaacttaatgctctcttggaatgt 4154
QY 181 gattcatctgggataaaacaacatatlttaacctccaattatgtcttcgaatccg 240
   |||||||
DB 4155 gattcatctgggataaaacaacatatlttaacctccaattatgtcttcgaatccg 4214
QY 241 ttggacacacacatcatatagacttcgacgactcttcgattgattgattggtgtga 300
   |||||||
DB 4215 ttggacacacacatcatatagacttcgacgactcttcgattgattgattggtgtga 4274
QY 301 tttaaatagttgacagatgcatttgggaatggagaaagacataatcagatcacagat 360
   |||||||
DB 4275 tttaaatagttgacagatgcatttgggaatggagaaagacataatcagatcacagat 4334
QY 361 tactgttactactacttactacaaatgtttgcacactgtctctcttcaaaagctcgat 420
   |||||||
DB 4335 tactgttactactacttactacaaatgtttgcacactgtctctcttcaaaagctcgat 4394
QY 421 tcaccccaaggaggaatgactcctggagacccctcagtggaataaccaaagagtgct 480
   |||||||
DB 4395 tcaccccaaggaggaatgactcctggagacccctcagtggaataaccaaagagtgct 4454
QY 481 gcaagtggacttcagaaacaaatgaaatcacaagagtaactactcagaggaataaact 540
   |||||||
DB 4455 gcaagtggacttcagaaacaaatgaaatcacaagagtaactactcagaggaataaact 4514
QY 541 tctgcttaccagatgattgtaagagttctctcatctccacagatcaagaatggccatca 600
   |||||||
DB 4515 tctgcttaccagatgattgtaagagttctctcatctccacagatcaagaatggccatca 4574
QY 601 gtagactctcttttcagaatgagcaaatgaaagttttcagggaataacgaactcctt 660
   |||||||
DB 4575 gtagactctcttttcagaatgagcaaatgaaagttttcagggaataacgaactcctt 4634
QY 661 cagactgtgggaactctctctacacacgcttactgactcgtacactcgaatcaccc 720
   |||||||
DB 4635 cagactgtgggaactctctctacacacgcttactgactcgtacactcgaatcaccc 4694
QY 721 ccagagttgggtgcacacacattgctcctgagatgaggttttcgggctcggggcaga 780
   |||||||
DB 4695 ccagagttgggtgcacacacattgctcctgagatgaggttttcgggctcggggcaga 4754
QY 781 cctctactgactcgagcctaataaagaattatcttcatgacatagtggtgtgtt 840
   |||||||
DB 4755 cctctactgactcgagcctaataaagaattatcttcatgacatagtggtgtgtt 4809
QY 841 ttgtgtgtggtgcggcagaagaacccctagtgatgaggttgccaactcctctgcgagctc 900
   |||||||
DB 4810 ttgtgtgtggtgcggcagaagaacccctagtgatgaggttgccaactcctctgcgagctc 4869
QY 901 gctcgtactacagcgcgcgacccaagatgcggcgacggcgcggtcttgccggggcg 960
   |||||||
DB 4870 gctcgtactacagcgcgcgacccaagatgcggcgacggcgcggtcttgccggggcg 4929
QY 961 cctcagtgagcgagcgagcgcgagctgctctgcaggaat 1000
   |||||||
DB 4930 cctcagtgagcgagcgagcgcgagctgctctgcaggaat 4969

```

```

RESULT 4
AAD08612
ID AAD08612 standard; DNA, 11933 BP.
XX
AC AAD08612;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human factor VIII expressing RAAV vector PAAV-F8-1 partial sequence.
XX
KW Recombinant adeno-associated viral vector; RAAV; factor VIII; FVIII;
KW blood clotting disorder; gene therapy; haemophilia A; human;
KW PAAV-F8-1; ds.
XX
OS Chimeric - Adeno associated virus.
OS Chimeric - Mus sp.
OS Chimeric - Homo sapiens.
OS Chimeric - Oryctolagus cuniculus.
XX
PN W0200145510-A1.
XX
PD 28-JUN-2001.
XX
PF 21-DEC-2000; 2000WO-US34925.
XX
PR 22-DEC-1999; 99US-0470618.
XX
PA (AVIG-) AVIGEN INC.
XX
PI Couto LB, Colosi PC, Qian X;
XX
PI WPI: 2001-417955/44.
XX
DR
PT Treating blood clotting disorder, especially hemophilia in mammals, by
PT administering recombinant adeno-associated vectors which express blood
PT coagulation factor VIII -
XX
PS Claim 18; Fig 5; 90pp; English.
XX
CC The present invention relates to a method for treating a subject
CC suffering from a blood clotting disorder. The method comprises
CC administering a recombinant adeno-associated virion (RAAV) comprising
CC a nucleotide sequence encoding the light chain of factor VIII and a
CC second recombinant adeno-associated virion comprising a nucleotide
CC sequence encoding the heavy chain of factor VIII. The RAAV vector is
CC useful in gene therapy for treating haemophilia A in mammals, in
CC particular humans. The RAAV vector provides high level and long term
CC expression of biologically active clotting factor VIII in vivo.
CC The present sequence is PAAV-F8-1 vector without the plasmid
CC backbone. It comprises HNF-3 mouse albumin promoter, a synthetic intron
CC based on human elongation factor Ialpha (EF1alpha) and immunoglobulin G
CC (196) intron sequences, B-domain deleted human factor VIII coding
CC sequence, poly A signal based on rabbit beta-globin sequence and AAV
CC inverted terminal repeats (ITRs) at the ends. The vector encodes both the
CC light and heavy chains of human factor VIII.
XX
SQ Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;

```

Query Match 95.4%; Score 954.4; DB 22; Length 11933;
 Best Local Similarity 97.9%; Pred. No. 1.7e-287;
 Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

```

QY 1 gatcaagtgatgctgttgacccaatgattatcacgacatcaagaaccagggtgccg 60
   |||||||
DB 3975 gatcaagtgatgctgttgacccaatgattatcacgacatcaagaaccagggtgccg 4034
QY 61 tcaaaattccacgactctacatctcagttatcatatcatatgattatgattggaa 120
   |||||||
DB 4035 tcgaagttccacgactctacatctcagttatcatatcatatgattatgattggaa 4094
QY 121 gaagtgacagactatcgaggaattccactggaaacttaatgctctcttggaatgt 180
   |||||||

```

Db	4095	gaagtgcgagacttaccgaggaattccacctggaaccttaatgtctctttgccaatgt	4154
Qy	181	ggatltcatctcgagatataaacacacaaatatttttaacctccaattatgtctgatacatcg	240
Db	4155	ggattcatctcgagatataaacacacaaatatttttaacctccaattatgtctgatacatcg	4214
Qy	241	tttgcaccaactcatattatagcatctgcgcagacctcttgcgaatgagtttgaaggcttga	300
Db	4215	tttgcaccaactcatattatagcatctgcgcagacctcttgcgaatgagtttgaaggcttga	4274
Qy	301	tttaaatagtgtgagcatgcatctggtgaaatgtagagatlaaagaacatatacatgacacgat	360
Db	4275	tttaaatagtgtgagcatgcatctggtgaaatgtagagatlaaagaacatatacatgacacgat	4334
Qy	361	tactgtctcaactcttaaccaatatgtttgccacctggtctctcttcaaaagctgact	420
Db	4335	tactgtctcaactcttaaccaatatgtttgccacctggtctctcttcaaaagctgact	4394
Qy	421	tcacctccaaggagagatgaatgctctggtgagacctgaagtgtgaataatccaagaagtgtgt	480
Db	4335	tcacctccaaggagagatgaatgctctggtgagacctgaagtgtgaataatccaagaagtgtgt	4454
Qy	481	gcaagtgagacttccagaagaacaatgtgaatgcacagagatgaactactcaaggagatfaaatc	540
Db	4455	gcaagtgagacttccagaagaacaatgtgaatgcacagagatgaactactcaaggagatfaaatc	4514
Qy	541	tctcttaccagaatgatatgtgaagaagttctccatctccagcaatcagatgagatgcacata	600
Db	4515	tctcttaccagaatgatatgtgaagaagttctccatctccagcaatcagatgagatgagatgcacata	4574
Qy	601	gtgagctctcttttttcaagaatggtcaaatgaaggtttttcagsgaaatcaagactcct	660
Db	4575	gtgagctctcttttttcaagaatggtcaaatgaaggtttttcagsgaaatcaagactcct	4634
Qy	661	cacacccgtgtgagacactctatagaaccaccgttactgagctgtctacattcgaaattcaccc	720
Db	4635	cacacccgtgtgagacactctctatagaaccaccgttactgagctgtctacattcgaaattcaccc	4694
Qy	721	ccaagatgtggtgtgacacagaattgcctctgaagatgtagaggttctgtgctgcgaagcaacaga	780
Db	4695	ccaagatgtggtgtgacacagaattgcctctgaagatgtagaggttctgtgctgcgaagcaacaga	4754
Qy	781	cctctactgagctgagcctataaagaatattatcttcatgtcaatatgattgtgtgtt	840
Db	4755	cctctactgagctgagcctataaagaatattatcttcatgtcaatatgattgtgtgtt	4805
Qy	841	tttgtgtgcgcgcgcagagaacccccctgaatgtgtagagttgtgcacatccctctctgcgcgttc	900
Db	4810	tttgtgtgcgcgcgcagagaacccccctgaatgtgtagagttgtgtgcacatccctctctgcgcgttc	4865
Qy	901	gtctgcactactgagccggcgacacaaagtgtgcccgagcccggtgtcttgcgcggcg	960
Db	4870	gtctgcactactgagccggcgacacaaagtgtgcccgagcccggtgtcttgcgcggcg	4925
Qy	961	cctcagtgagcgagcagacgcgcagcgtcgtcctgcaaggaat	1000
Db	4930	cctcagtgagcgagcagacgcgcagcgtcgtcctgcaaggaat	4969
RESULT 5			
AAAF84647			
ID AAF84647 standard; DNA; 7944 BP.			
AAAF84647;			
29-JUN-2001 (first entry)			
Plasmid DLZ6 encoding human B-domain deleted factor VIII.			
Adeno-associated virus vector: B-domain; factor VIII; haemophilic A;			
coagulation disorder; ss.			

OS	Synthetic.
OS	Homo sapiens.
OS	Hepatitis B virus.
XX	
FH	Key
FT	Location/Qualifiers
FT	misc.feature
FT	I..146
FT	/tag= a
FT	/note= "Inverted terminal repeat"
FT	enhancer
FT	150..278
FT	/tag= b
FT	/note= "hepatitis B virus EnhI enhancer"
FT	CDS
FT	420..4835
FT	/tag= c
FT	/note= "human B-domain deleted factor VIII"
FT	polyA_signal
FT	4840..4914
FT	/tag= d
FT	/note= "TK polyA sequence"
FT	misc.feature
FT	4916..5084
FT	/tag= e
FT	/note= "Inverted terminal repeat"
PN	
PD	W0200127303-A1.
PD	19-APR-2001.
XX	
PF	12-OCT-2000; 2000WO-US2822L.
XX	
PR	12-OCT-1999; 99US-0158780.
XX	
PA	(UNNC-) UNIV NORTH CAROLINA.
PI	
DR	Walsh CE, Chao H, Burstein H, Lynch CM, Stepan AM, Munson K;
DR	WPI; 2001-273781/28.
P-PSDB;	AAB67959.
XX	
PT	New recombinant adeno-associated virus vector, useful for treating
PT	haemophilia A, comprises heterologous nucleotide sequence encoding
PT	B-domain deleted human factor VIII operably linked with liver-preferred
PT	expression control element -
PS	
XX	Claim 64; Fig 1; 87pp; English.
XX	
CC	The specification describes a recombinant adeno-associated virus (RAV)
CC	vector. The vector comprises a heterologous nucleotide sequence
CC	encoding B domain deleted factor VIII operably linked with at least one
CC	enhancer and at least one promoter. The method results in the production
CC	of high titer RAV vector stocks carrying the B-domain deleted factor
CC	VIII transgenes and expression cassettes, which generate adequate titers
CC	of viruses for in vivo administration. The recombinant vectors are useful
CC	for treating haemophilia A, where the liver expresses the encoded
CC	B-domain deleted factor VIII, which is secreted into the blood. They are
CC	also useful for the treatment of other coagulation disorders. The
CC	present sequence encodes a B-domain deleted factor VIII.
XX	
SQ	Sequence 7944 BP: 2142 A; 1902 C; 1909 G; 1991 T; 0 other:
	Query Match 87.1%; Score 870.8; DB 22; Length 7944;
	Best Local Similarity 92.7%; Pred.No.1.8e-261;
	Matches 933; Conservative 0; Mismatches 52; Indels 21; Gaps 12;
OY	1 gatcaaggatgacgtctgtggacaacatgatatttcacggcatcacaagaaccagagtcccg 60
DG	
DG	4046 gatcaaggatgacgtctgtggacaacatgatatttcacggcatcacaagaaccagagtcccg 4105
OY	61 tcagaagttctccagccttaaatctctcaagtttatcatcatcatgataagcttgatgggaa 120
DG	
DG	4106 tcagaagttctccagccttaaatctctcaagtttatcatcatcatgataagcttgatgggaa 4165
OY	121 gaagtggacgaactatccggggaattccaactgaaccttaagtctctcttggcaatgt 180
DG	
DG	4166 gaagtggacgaactatccggggaattccaactgaaccttaagtctctcttggcaatgt 4225

[illegible]

RESULT	6
AAAN90654	
ID	AAAN90654 standard; DNA; 4275 BP.
XX	
AC	AAAN90654;
XX	
DT	26-JUN-1990 (first entry)
XX	
DE	DNA encoding 740 Arg-1649 Glu human Factor VIII:C.
XX	
KW	Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;
KW	haemophilia A.
XX	
OS	Homo sapiens.
XX	

Key	Location/Qualifiers
CD5	1..4275
	/+tag= a
EP306968-A.	
15-MAR-1989.	
09-SEP-1988;	88EP-0114769.
08-APR-1988;	88JP-0085454.
(KAGA)	CHEMO-SERO-THERAP.
(TEIJ)	TEIJIN LTD.
Sugiyama T, Masuda K, Tajima Y, Yonemura H;	
WPI; 1989-078467/11.	
P-PSDB; AAP91165.	
Prodn. of recombinant human Factor-VIII-C -	
using animal cells transformed with a vector contg. the gene for	
Factor VIII:C and a promoter	
Fig 1(1) - 1(13); ; 32pp; English.	
When translated, Arg-740 of the carboxyl terminus of the H chain is	
directly bonded by a peptide bond to Glu-1649 of the amino terminus of	
L chain. It is used to transform animal cells so that they produce	
human Factor VIII:C. A pred. expression vector is plasmid Ad.RE.neo.	
The expression vector has at least one promoter upstream of AAN90654.	
The transformants can constantly and continuously produce human Factor	
VIII:C in high yield on a commercial scale. The human Factor VIII:C so	
produced is considered to corresp. to the smallest species of active and	
intact Factor VIII:C molecules in the human blood plasma. It is useful	
for treating haemophilia A patients.	
Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;	

Query Match	Similarity	79.0%	Score 790	DB 10	Length 4275
Best Local	Similarity	100.0%	Pred. NO. 2,4e+26		
Matches	790	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY 1	gatacaagtgatgctgtttggccaacatgatlttaacggcatcaagaaccaggtgccc		60		
Db 3486	gatacaagtgatgctgtttggccaacatgatlttaacggcatcaagaaccaggtgccc		3545		
QY 61	tcaagaattcttcaacccctacatctcttcacgtttatcatcatgatatagtcttgatggaa		120		
Db 3546	tcaagaattcttcaacccctacatctcttcacgtttatcatcatgatatagtcttgatggaa		3605		
QY 121	gaagtgagacactlctgaggaatctccacgttgaaaccttaatgtgctctcttgacaagt		180		
Db 3606	gaagtgagacactlctgaggaatctccacgttgaaaccttaatgtgctctcttgacaagt		3665		
QY 181	ggattccatctgggataaaaacacataatttttaacctccaattatgctcgatataccg		240		
Db 3666	ggattccatctgggataaaaacacataatttttaacctccaattatgctcgatataccg		3725		
QY 241	tttgacccaactcactlaataagatctgcagcacactcttcgcagtggatgtgagggctgga		300		
Db 3726	tttgacccaactcactlaataagatctgcagcacactcttcgcagtggatgtgagggctgga		3785		
QY 301	tttaatatgtttgcagcatgccaattgaggaaatgagataaagaacatatacatgacacagat		360		
Db 3786	tttaatatgtttgcagcatgccaattgaggaaatgagataaagaacatatacatgacacagat		3845		
QY 361	tacgcttcacacacttaccataatgtttggccaacctggtctctcttcaaaaagctgaat		420		
Db 3846	tacgcttcacacacttaccataatgtttggccaacctggtctctcttcaaaaagctgaat		3905		
QY 421	tcacctccaaggagagtaatgctctgagacactcaagttgataatccaanaagttgct		480		

```

|||||
Db 3906 tcacctcaaggagagtaatgctcctgagaccccaagtgataatcccaaaagatggct 3965
Qy 481 gcaagtggactccagaagaacatgaagtcacagagtaactactcaaggagtaaatc 540
Db 3966 gcaagtggactccagaagaacatgaagtcacagagtaactactcaaggagtaaatc 4025
Qy 541 tctgttaccagatgtatgtgaaggagttccatccatccagcagtaagaatggcactca 600
Db 4026 tctgttaccagatgtatgtgaaggagttccatccatccagcagtaagaatggcactca 4085
Qy 601 gtggactctcttttccagaatggcaagtaagttttcagggaatcaagactcctt 660
Db 4086 gtggactctcttcttcagaatggcaagtaagttttcagggaatcaagactcctt 4145
Qy 661 cacacctgtgtgaactctctcagaccacgcttactgactcgtactccttcgaattcacc 720
Db 4146 cacacctgtgtgaactctctcagaccacgcttactgactcgtactccttcgaattcacc 4205
Qy 721 ccagagttgggtgcacacagatgcctcgaagatggagttctggcctgcagagcacaga 780
Db 4206 ccagagttgggtgcacacagatgcctcgaagatggagttctggcctgcagagcacaga 4265
Qy 781 cctctactga 790
Db 4266 cctctactga 4275

RESULT 7
AAV23339
ID AAV23339 standard; DNA: 4629 BP.
AC AAV23339;
XX
XX 24-SEP-1999 (first entry)
XX
XX Human Factor VIII with B domain deleted cDNA.
XX
XX Adenoviral vector; Factor VIII; Factor IX; clotting factor; treatment;
XX haemostatic; haemophilia A; haemophilia B; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX US5935935-A.
XX
XX 10-AUG-1999.
XX
XX 07-JUN-1995; 95US-0484891.
XX
XX 07-JUN-1995; 95US-0484891.
XX
XX 10-JUN-1993; 93US-0074920.
XX
XX 25-MAR-1994; 94US-0218335.
XX
XX (GENE-) GENETIC THERAPY INC.
XX
XX Connelly S, Kaleko M, Smith T;
XX
XX WPI; 1999-457617/38.
XX
XX Adenoviral vectors useful for treating hemophilia
XX
XX Example 1; Column 53-58; 90pp; English.
XX
XX This invention describes novel adenoviral vectors comprising at least
XX one DNA sequence encoding a clotting factor (Factor IX or Factor VIII).
XX The vectors of the invention have haemostatic activity. The vectors are
XX useful for the treatment of hemophilia A or hemophilia B by gene therapy.
XX This sequence represents human Factor VIII cDNA which has the B domain
XX deleted.
XX
XX Sequence 4629 BP; 1319 A; 1050 C; 1015 G; 1245 T; 0 other;
XX

```

```

Query Match 79.0%; Score 790; DB 20; Length 4629;
Best Local Similarity 99.4%; Pred. No. 2,5e-236;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 gatcaagttgatctgttgcacaaatgatattatccagcatcaaaagagtgccg 60
Db 3627 gatcaagttgatctgttgcacaaatgatattatccagcatcaaaagagtgccg 3686
Qy 61 tcgaagttctccagctcctcactcctcagtttatccatcatatagcttgatggaa 120
Db 3687 tcgaagttctccagctcctcactcctcagtttatccatcatatagcttgatggaa 3746
Qy 121 gaagtgcagactatccagaagaatccactggaaacttaatgttctcttggcaatg 180
Db 3747 gaagtgcagactatccagaagaatccactggaaacttaatgttctcttggcaatg 3806
Qy 181 ggaattcatctggataaacaacatatatttaacctccaattatgtctgcgaactcg 240
Db 3807 ggaattcatctggataaacaacatatatttaacctccaattatgtctgcgaactcg 3866
Qy 241 ttggacacacactatcatatagcattcgcagacactcttcgcatggattggtgctgta 300
Db 3867 ttggacacacactatcatatagcattcgcagacactcttcgcatggattggtgctgta 3926
Qy 301 tttaaatgattgcagcatgcatatggagaaatggagaaatgaacataatcagatgacagat 360
Db 3927 tttaaatgattgcagcatgcatatggagaaatggagaaatgaacataatcagatgacagat 3986
Qy 361 tactgtctcatcctactcttaccataatgtttgcacactgtgtcccttcaaaactgcact 420
Db 3987 tactgtctcatcctactcttaccataatgtttgcacactgtgtcccttcaaaactgcact 4046
Qy 421 tcacctcaaggagaggtatgtcctcgtgagacactcagtgtaataatccaaagatggct 480
Db 4047 tcacctcaaggagaggtatgtcctcgtgagacactcagtgtaataatccaaagatggct 4106
Qy 481 gcaagtggactccagaagaacatgaagtcacagagtaactactcaaggagtaaatc 540
Db 4107 gcaagtggactccagaagaacatgaagtcacagagtaactactcaaggagtaaatc 4166
Qy 541 tctgttaccagatgtatgtgaaggagttccatccatccagcagtaagaatggcactca 600
Db 4167 tctgttaccagatgtatgtgaaggagttccatccatccagcagtaagaatggcactca 4226
Qy 601 gtggactctcttttccagaatggcaagtaagttttcagggaatcaagactcctt 660
Db 4227 gtggactctcttcttcagaatggcaagtaagttttcagggaatcaagactcctt 4286
Qy 661 cacacctgtgtgaactctctcagaccacgcttactgactcgtactccttcgaattcacc 720
Db 4287 cacacctgtgtgaactctctcagaccacgcttactgactcgtactccttcgaattcacc 4346
Qy 721 ccagagttgggtgcacacagatgcctcgaagatggagttctggcctgcagagcacaga 780
Db 4347 ccagagttgggtgcacacagatgcctcgaagatggagttctggcctgcagagcacaga 4406
Qy 781 cctctactcagagcc 798
Db 4407 cctctactcagaggtggcc 4424

RESULT 8
AAV23339
ID AAV23339 standard; DNA: 4670 BP.
AC AAV23339;
XX
XX 17-AUG-1998 (first entry).
XX
XX Human Factor-VIII gene lacking central B domain.
XX
XX Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.
XX

```

OS Homo sapiens.
OS Synthetic.
XX
XX
PN WO9812207-A1.
XX
XX
PD 26-MAR-1998.
XX
XX
PF 18-SEP-1997; 97WO-US16639.
XX
PR 20-SEP-1996; 96US-0717294.
XX
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
XX
PI Haas J, Seed B;
XX
XX
DR WPI; 1998-217200/19.
XX
XX
PT New synthetic eukaryotic gene(s) - in which non-preferred or less
PT preferred codon(s) are replaced to provide high level expression in
XX mammalian cell(s)

Example 3; Fig 12; 92pp; English.

CC This gene codes for a human Factor-VIII protein that lacks the
CC central B domain (amino acids 760-1630) of the native protein. In
CC a novel, claimed synthetic gene (see AAV32388), non-preferred or
CC less preferred codons of the native gene are replaced by codons
CC favored by highly expressed human genes to provide high level
CC expression in mammalian cells. The synthetic gene was assembled
CC from 29 pairs of oligonucleotides* (see AAV23340-97) which served as
CC PCR templates. Synthetic genes of the invention (see also
CC AAV23289-91) are used for production of recombinant proteins in
CC mammalian cells at levels of at least 500% of those obtained using
CC the natural genes. They can also be used in gene therapy.

SQ Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T; 0 other;

Query Match	79.0%;	Score 790;	DB 19;	Length 4670;
Best Local Similarity	99.4%;	Pred. No. 2.5e-236;		
Matches 793;	Conservative	0;	Mismatches 5;	Indels 0;
				Gaps 0;

OY	1	gatcaagtgatctgtttggaaccaatgatatttcacgagatcaaaacccagggtgccg	60
Db	3653	gatcaagtgatctgtttggtgcaccaatgatatttcacgagatcaaaacccagggtgccg	3712
OY	61	tcagaagttctccagcctctacatctctccagttatcatcatgtatagcttgatggaa	120
Db	3713	tcagaagttctccagcctctacatctctccagttatcatcatgtatagcttgatggaa	3772
OY	121	gaatgtggaagcttatgtgaagaaattccactgtgaacctaaatggtctctttggcaagt	180
Db	3773	gaatgtggaagcttatgtgaagaaattccactgtgaacctaaatggtctctttggcaagt	3832
OY	181	ggattcctctgggataaacaacataatttttaacctccaattatgtctcgatataccg	240
Db	3833	ggattcctctgggataaacaacataatttttaacctccaattatgtctcgatataccg	3892
OY	241	tttgcacccaactcatatatgatcatcttcgcagcactcttcgcatagtgttatgtgctgtga	300
Db	3893	tttgcacccaactcatatatgatcatcttcgcagcactcttcgcatagtgttatgtgctgtga	3952
OY	301	tttaaacagtctgcagcatgtgcattctggtaatgtgaagttaaagaacataatgaatgcacgat	360
Db	3953	tttaaacagtctgcagcatgtgcattctggtaatgtgaagttaaagaacataatgaatgcacgat	4012
OY	361	tactgcttcaactcaacttacaacatatgtttgcacactgtgtctctctcaaaagctcgact	420
Db	4013	tactgcttcaactcaacttacaacatatgtttgcacactgtgtctctctcaaaagctcgact	4072
OY	421	tcacctccaaggaggagtaatgtccctggagacctgaagtgtgaataatccaagaagtgctc	480
Db	4073	tcacctccaaggaggagtaatgtccctggagacctgaagtgtgaataatccaagaagtgctc	4132

OY	481	gcaagtgactccaagaagcaatgtgaagtcacaagaagtaactactccagggtataatc	540
Db	4133	gcaagtgactccaagaagcaatgtgaagtcacaagaagtaactactccagggtataatc	4132
OY	541	tcgtctaccagcatgtatgtgaaggagttcctcatalcctccagcagtcagaatgscatca	600
Db	4193	tcgtctaccagcatgtatgtgaaggagttcctcatalcctccagcagtcagaatgscatca	4252
OY	601	gtgactctctttttccaagatgccaagttaaagttcttcagggaataccaagactcctt	660
Db	4253	gtgactctctttttccaagatgccaagttaaagttcttcagggaataccaagactcctt	4312
OY	661	caaccctgtgtgaactctcttagaccaccgttactgaactgtctactctgaattaccc	720
Db	4313	caaccctgtgtgaactctcttagaccaccgttactgaactgtctactctgaattaccc	4372
OY	721	ccaagattgtgtgcacacagaatttcctcttgagatggaatggaattctcgggctcgagcacaga	780
Db	4373	ccaagattgtgtgcacacagaatttcctcttgagatggaatggaattctcgggctcgagcacaga	4432
OY	781	ccttactgaactgaacc	798
Db	4433	ccttactgaactgaacc	4450

RESULT

ID AAV19581 standard; cDNA; 4832 BP.

AC AAV19581;

DT 06-AUG-1998 (first entry)

Human factor VIII beta-domain deleted SQN deletion cDNA sequence.

KM Replication defective; recombinant retrovirus: RV; therapeutic protein;
KM haemophilia; thrombosis; hypercoagulable disorder: liver disease; human;
KM hepatitis; thalassemia; phenylketonuria; Leach-Nyman syndrome; diabetes;
KM cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;
KM hypophosphatemia; adenine deaminase deficiency; HIV infection; anaemia;
KM Gaucher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;
KM inflammatory disease; factor VIII:ss.

05 Homo sapiens.

PN W09800541-A2.

PD 08-JAN-1998.

PF 02-JUL-1997; 97WO-US11784.

PR 04-JUN-1997; 97US-0869309.

PR 13-AUG-1996; 96US-0696381.

PA (CHIR) CHIRON CORP.

PI Allen JR, Barber JR, Boder M, Chang SMW, Chong K;

PI Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respress J

WPI; 1998-086966/08.

XX
XX
XX

PT administered to provide long term systemic expression of therapeutic

PT disorders

Example 28; Pages 210-213; 272pp; English.

CC This cDNA encodes the beta-domain deleted SQN deletion protein of human

CC factor VIII. This is used in the construction of recombinant retroviral
 CC vectors expressing human factor VIII. The invention provides the
 CC preparation of replication defective recombinant retrovirus (RV)
 CC expressing a therapeutic protein. The RV preparation is resistant to
 CC degradation by human complement and is capable of inducing long term
 CC systemic expression of the therapeutic protein when administered
 CC intravenously to a human. The long term systemic expression results in a
 CC measurable level of the therapeutic protein being produced in the blood
 CC of the human for a period of at least 30 days after the administration of
 CC the RV vector preparation. RV's can be used for in vivo delivery of
 CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,
 CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,
 CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,
 CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's
 CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,
 CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha-
 CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as
 CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or
 CC inflammatory disease or graft versus host disease. RV's are capable of
 CC surviving inactivation in human serum thereby allowing efficient gene
 CC transfer over prolonged periods of time.

XX Sequence 4832 BP, 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 79.0%; Score 790; DB 19; Length 4832;
 Best Local Similarity 99.4%; Pred. No. 2, 6e-236;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaggtgatctgttgcacccaatgattatcaagcagatcaagccagggtgccg 60
 Db 3656 gataaggtgatctgttgcacccaatgattatcaagcagatcaagccagggtgccg 3715
 QY 61 tcagaagttctcagcctctacatctctcagttatcatatgattgttttgagaa 120
 Db 3716 tcagaagttctcagcctctacatctctcagttatcatatgattgttttgagaa 3775
 QY 121 gaattgcagactatcgagaatttcactgcagacttaattgtctctcttgcaatt 180
 Db 3776 gaattgcagactatcgagaatttcactgcagacttaattgtctctcttgcaatt 3835
 QY 181 gattcatcttgcgttaaaacacaaatattttaaaccctcaattatgtctgcataacacg 240
 Db 3836 gattcatcttgcgttaaaacacaaatattttaaaccctcaattatgtctgcataacacg 3895
 QY 241 ttctgcacccaactattatagatctgcagcactcttcgatgagttgtgtgtgtga 300
 Db 3896 ttctgcacccaactattatagatctgcagcactcttcgatgagttgtgtgtgtga 3955
 QY 301 ttttaattgttcagcattgcattgggaattggaattgaagaatcaatcaatgcacagat 360
 Db 3956 ttttaattgttcagcattgcattgggaattggaattgaagaatcaatcaatgcacagat 4015
 QY 361 tactgtctcatctacttaccatattgttgcacactggtctctctcaaaagctgact 420
 Db 4016 tactgtctcatctacttaccatattgttgcacactggtctctctcaaaagctgact 4075
 QY 421 tcaactcagaaggaggaatgctcgtgagacctcagtgatgataatccaaagggtgct 480
 Db 4076 tcaactcagaaggaggaatgctcgtgagacctcagtgatgataatccaaagggtgct 4135
 QY 481 gcaattgacttcagaagacatgaagtcacaggaattactactcagggaggaatc 540
 Db 4136 gcaattgacttcagaagacatgaagtcacaggaattactactcagggaggaatc 4195
 QY 541 tctgtttaccagcagatgtatgtgaaggagttcctcatctccagcagtcgaagtgccatca 600
 Db 4196 tctgtttaccagcagatgtatgtgaaggagttcctcatctccagcagtcgaagtgccatca 4255
 QY 601 gtgagactcttttttcagaatgagcaagtaagtttttaagggaatcagactcctt 660
 Db 4256 gtgagactcttttttcagaatgagcaagtaagtttttaagggaatcagactcctt 4315

QY 661 cacactgtgtgactctctagaccacacgcttactgactcgtactcgaattacacc 720
 Db 4316 cacactgtgtgactctctagaccacacgcttactgactcgtactcgaattacacc 4375
 QY 721 ccagagttgggtgcacacagattcccttgagatggaattcttgggtcgcagacagga 780
 Db 4376 ccagagttgggtgcacacagattcccttgagatggaattcttgggtcgcagacagga 4435
 QY 781 cctctactgactcagacc 798
 Db 4436 cctctactgaggtgtgcc 4453

RESULT 10
 ID AAV15338 standard; DNA; 4832 BP.
 AAV15338;
 AAV15338;

XX 20-JUL-1998 (first entry)

DE Human Factor VIII SON deletion mutant DNA.

KW Factor VIII; blood clotting; haemophilia A; gene therapy;
 KW retrovirus; vector; human; ss.

OS Homo sapiens.
 OS Synthetic.

XX Key Location/Qualifiers
 XX CDS 72..4445
 XX /*tag= a

PN MO9800542-A2.

PD 08-JAN-1998.

XX 02-JUL-1997; 97WO-US11785.

XX 04-JUN-1997; 97US-0869309.

PR 03-JUL-1996; 96US-0645601.

PR 13-AUG-1996; 96US-0696381.

XX (CHIR) CHIRON CORP.

PI Allen JR, Barber JR, Boder M, Chang SM, Chong K;
 PI De la Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;
 PI Jolly DJ, Mittelstaedt DM, Prussak CE, Respass JG;

XX WPI: 1998-086967/08.
 XX P-PSDB: AAW4372.

PT New replication defective recombinant retroviruses - which express B
 PT domain-deleted human factor VIII or human factor IX for the

PT treatment of haemophilia

XX Claim 6; Page 174-175; 236pp; English.

XX This DNA sequence includes a coding region for the B domain
 CC deletion mutant SON (see AAW4372) of human Factor VIII. The SON
 CC mutant is created by fusing Ser-743 to Gln-1638 of native Factor
 CC VIII (see AAW4373) to form a Ser-Gln-Asn (SON) link between the A2
 CC and A3 Factor VIII domains. When compared to plasmid-derived
 CC Factor VIII, the SON deletion does not influence the in vivo
 CC pharmacokinetics, but the reduced size of the molecule appears to
 CC decrease proteolytic degradation. The invention relates to
 CC preparations of replication defective recombinant retrovirus (RV)
 CC expressing a B domain-deleted human factor VIII protein, where the
 CC recombinant RV is capable of infecting human cells, is resistant
 CC to degradation by human complement and is capable of inducing
 CC long-term (at least 30 days and up to 6 months or longer
 CC post-injection) systemic expression of Factor VIII when
 CC administered to a haemophilia A patient.

XX Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 79.0%; Score 790; DB 19; Length 4832;
Best Local Similarity 99.4%; Pred. No. 2.6e-236;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 gatcaaggtgagctgttggcaccatgattatcaaggatcaaacaccagggtgccg 60
   |||
DB 3656 gatcaaggtgagctgttggcaccatgattatcaaggatcaaacaccagggtgccg 3715
QY 61 tcagaagttctcagcctctacatctcagttatcatcatgatagtctgatggaa 120
   |||
DB 3716 tcagaagttctcagcctctacatctcagttatcatcatgatagtctgatggaa 3775
QY 121 gaagtgagagactatcagaggaattccatcgaaacttaagtgcttcttggcaatg 180
   |||
DB 3776 gaagtgagagactatcagaggaattccatcgaaacttaagtgcttcttggcaatg 3835
QY 181 ggaattacatgagataaacaacataatttaacctccattatgtctgatacatcgg 240
   |||
DB 3836 ggaattacatgagataaacaacataatttaacctccattatgtctgatacatcgg 3895
QY 241 ttgacccaactcatatataagcattcgacacctctcgatggagttgagtggtgtga 300
   |||
DB 3896 ttgacccaactcatatataagcattcgacacctctcgatggagttgagtggtgtga 3955
QY 301 tttaaatgttcagacatgccattgggaatggagtaagaacatatacagatgccagat 360
   |||
DB 3956 tttaaatgttcagacatgccattgggaatggagtaagaacatatacagatgccagat 4015
QY 361 tactgtctatcctcatcttaccataatgtttgcacccgtgtctcttcaaaaagctcgat 420
   |||
DB 4016 tactgtctatcctcatcttaccataatgtttgcacccgtgtctcttcaaaaagctcgat 4075
QY 421 taactctcaaggaggagtaatgtctgtgagacctcagggtgaataatccaaaagagtgct 480
   |||
DB 4076 taactctcaaggaggagtaatgtctgtgagacctcagggtgaataatccaaaagagtgct 4135
QY 481 gaaatgtgacctccagaaacaataaagtcacagagtaactactactcaggagtaaatc 540
   |||
DB 4136 gaaatgtgacctccagaaacaataaagtcacagagtaactactactcaggagtaaatc 4195
QY 541 tctgttaccagacatgatagttgaaggagttcctcatctccagcagtcagaagtggcatca 600
   |||
DB 4196 tctgttaccagacatgatagttgaaggagttcctcatctccagcagtcagaagtggcatca 4255
QY 601 gtagagctctcttttcagaatgagcaagtaaggttttcaagggaatcaagactcct 660
   |||
DB 4256 gtagagctctcttttcagaatgagcaagtaaggttttcaagggaatcaagactcct 4315
QY 661 cacacctgtgtgaactctctagaccacgcttactgactgctacacttgaaatccacc 720
   |||
DB 4316 cacacctgtgtgaactctctagaccacgcttactgactgctacacttgaaatccacc 4375
QY 721 ccagaagtgtgagtcacagacatgacctgaggaatgaggttctggcgtgcagagcaagga 780
   |||
DB 4376 ccagaagtgtgagtcacagacatgacctgaggaatgaggttctggcgtgcagagcaagga 4435
QY 781 cctctactgactgagcc 798
   |||
DB 4436 cctctactgaggtggcc 4453

```

RESULT 11

ID AAT69811 standard; DNA: 5035 BP.

XX AAT69811;

XX 10-AUG-1997 (first entry)

DE Factor VIII-dB695-HCII DNA.

KM Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
KW blood clotting; procoagulant; anticoagulant; antithrombotic;
KW haemophilia; gene therapy; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS 35..5020

FT misc_RNA 2225..2314

FT /product= Factor VIII-dB695-HCII
FT /tag= b
FT /product= heparin cofactor II region (aa51-81)

W09718315-A1.

22-MAY-1997.

PE 13-NOV-1996; 96WO-EP04977.

PR 13-NOV-1995; 95US-0558107.

PA (IMMO) IMMUNO AG.

PI Voorberg JJ;

DR WPI; 1997-289291/26.

DR P-PSDB; AAW18670.

PT Hybrid Factor VIII with modified activity, comprises region from
PT donor anticoagulant or antithrombotic protein - useful for treatment
PT of coagulation disorders

PS Claim 16; Page 52-60; 96pp; English.

CC A DNA molecule (AAT69811) codes for Factor VIII-dB695-HCII (AAW18670),
CC a hybrid protein in which amino acids 712-736 of Factor-dB695

CC (Factor VIII del868-1562) B-domain are replaced by amino acids

CC 51-80 from the acidic region (and potential thrombin-binding site)

CC of human heparin cofactor II (HCII). It was obtd. by PCR

CC amplification (see also AAT69812-13) of the HCII acidic region from

CC total liver cDNA, fusion to sequences encoding Factor VIII

CC aa706-711 and aa737-743, and incorporation of the construct into

CC plasmid pCUB-dB695. The hybrid protein, which can be expressed

CC using gene therapy techniques, has increased procoagulant activity

CC owing to the HCII acidic region, and can be used to treat blood

CC coagulation disorders such as haemophilia A.

XX Sequence 5035 BP; 1484 A; 1127 C; 1110 G; 1314 T; 0 other;

Query Match 79.0%; Score 790; DB 18; Length 5035;

Best Local Similarity 99.4%; Pred. No. 2.6e-236;

Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 gatcaaggtgagctgttggcaccatgattatcaaggatcaaacaccagggtgccg 60
   |||
DB 4231 gatcaaggtgagctgttggcaccatgattatcaaggatcaaacaccagggtgccg 4290
QY 61 tcagaagttctcagcctctacatctcagttatcatcatgatagtctgatggaa 120
   |||
DB 4291 tcagaagttctcagcctctacatctcagttatcatcatgatagtctgatggaa 4350
QY 121 gaagtgagagactatcagaggaattccatcgaaacttaagtgcttcttggcaatg 180
   |||
DB 4351 gaagtgagagactatcagaggaattccatcgaaacttaagtgcttcttggcaatg 4410
QY 181 ggaattacatgagataaacaacataatttaacctccattatgtctgatacatcgg 240
   |||
DB 4411 ggaattacatgagataaacaacataatttaacctccattatgtctgatacatcgg 4470

```


QY	241	tttgcacccaactcaatataagattatgcgaagacactcttcgcatlbgaaatttgatggccttga	300
QY	4471	tttgcgccaacacccattatagatctcgaagacactcttcgcattggaatttgatggccttga	4530
QY	301	tttaaatagtgtgacgatgccttgggaatggaatgaagaacaataloagaattgcacagat	360
Db	4531	tttaaatagtgtgacgatgccttgggaatggaatggaatgaagaacaataloagaattgcacagat	4590
QY	361	tactgcttcaactactattaccaataatgtttgcacactggtctcctcttcaaaagctgact	420
Db	4591	tactgcttcaactactattaccaataatgtttgcacactggtctcctcttcaaaagctgact	4650
QY	421	tcaccttccaaggaggaagaaatgacctggcgggagacttaaggttgaataatccaagaagtgct	480
Db	4651	tcaccttccaaggaggaagaaatgacctggcgggagacttaaggttgaataatccaagaagtgct	4710
QY	481	gcaagtgcattccagaagaacaatagaanaatcacaaggatlaactatccaaggagataaatc	540
Db	4711	gcaagtgcattccagaagaacaatagaanaatcacaaggatlaactatccaaggagataaatc	4770
QY	541	tctgtctaccagacatgtatgttgaaaggagttcctcatctccagcagtcgaagtggccatca	600
Db	4771	tctgtctaccagacatgtatgttgaaaggagttcctcatctccagcagtcgaagtggccatca	4830
QY	601	gtggaactcttttttccaagaatgagcaaatgaagtttttcaagggaataccaagactcctt	660
Db	4831	gtggaactcttttttccaagaatgagcaaatgaagtttttcaagggaataccaagactcctt	4890
QY	661	cacacccgtgtgtaaacctctctagaaccacacgcttactgactgcctactccttgaatcaccc	720
Db	4891	cacacccgtgtgtaaacctctctagaaccacacgcttactgactgcctactccttgaatcaccc	4950
QY	721	ccagagttgggtgcacacagatctgccttgaaagatggaaggtcttcgggctgcggagcaacaga	780
Db	4951	ccagagttgggtgcacacagatctgccttgaaagatggaaggtcttcgggctgcggagcaacaga	5010
QY	781	gccttactgactgcagcc 798	
Db	5011	ccttactatgaggtgcgc 5028	

RESULT	12
AAAA9231	
ID	AAAA9231 standard; DNA: 5094 BP.
XX	
AC	AAAA9231;
XX	
DT	26-SEP-2000 (first entry)
XX	
DE	DNA construct HSQ/EGFP for transforming endothelial cells.
XX	
KW	Hemostatic; antianemic; antidiabetic; nootropic; neuroprotective;
KW	osteopathic; antisticking; immunostimulant; gene therapy; collagen;
KW	endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;
KW	vascular endothelial growth factor; bovine brain extract; hemophilia;
KW	Factor VIII; human; transgene; adenosine deaminase deficiency; ss;
KW	sickle cell anaemia; thalassemia; diabetes; alpha-antitrypsin deficiency;
KW	Alzheimer's disease; brain disease; heart disease; immune system defect;
KW	bone fracture; osteoporosis.
XX	
OS	Homo sapiens.
XX	
PN	WO200032750-A1.
XX	
PD	08-JUN-2000.
XX	
PF	24-NOV-1999; 99WO-US28033.
XX	
PR	24-NOV-1998; 98US-0109687.
XX	
PA	(MIND) UNIV MINNESOTA.
PA	(UYEN-) UNIV EMORY.
PA	(HEBB/) HEBBEL R P.

PA (LINY/) LIN Y.
PA (LOLL/) LOLLAR J S.
XX
XX
P1 Hebbel RP, Lin Y, Lollar JS;
XX
DR WPI; 2000-412303/35.
XX
XX
XX Expanding population of endothelial cells useful to biocompatibilize
PT implantable medical devices comprises contacting buffy coat cells with
PT collagen I coated surface in culture medium comprising vascular
PT endothelial growth factor -
XX
XX
PS Claim 19; Fig 3; 53pp; English.

The invention relates to a method for expanding the population of endothelial cells (EC) obtained from peripheral blood, by culturing, in contact with a collagen I coated surface, buffy coat cells obtained from peripheral mammalian blood in the presence of a culture medium containing vascular endothelial growth factor (VEGF) and free of bovine brain extract. EC are useful for treating hemophilia by introducing it into the blood stream of a mammal, so that an effective amount of Factor VIII protein is secreted in the blood stream of the mammal. This sequence represents a human factor VIII gene into which an enhanced green fluorescent protein coding sequence has been inserted. The cDNA codes for (in order): the human Factor VIII (hVIII) activation peptide; the hVIII A1 domain; the hVIII A2 domain; the first 5 amino acids of the SQ B domain linker peptide; an enhanced green fluorescent protein (EGFP); the last 9 amino acids of the SQ linker peptide; the hVIII light chain activation peptide; the hVIII A3 domain; the hVIII C1 domain and the hVIII C2 domain. Transgenic EC transduced in vitro are useful for improving prosthetic implants. EC is also useful for diagnosing clotting disorders where indication of disease is associated with a reduction in the activity of an enzyme. EC is also useful in gene therapy for treating the variety of diseases, including adenosine deaminase deficiency, sickle cell anemia, thalassemia, hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders such as Alzheimer's disease, heart diseases, defects in immune system, for repairing bone fractures and to treat or prevent osteoporosis.

Query Match	79.0%	Score 790;	DB 21;	Length 5094;
Best Local Similarity	100.0%;	Fred. NO. 2.7e-226;		
Matches 790;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	gaccaagtgagatcgtgttgaccacaatgattatttcacggtacccaagaccaggtgcccg	60		
Db 4305	gaccaagtgagatcgtgttgaccacaatgattatttcacggtacccaagaccaggtgcccg	4364		
QY 61	tcagaagtttcacgacctacatctctcaagttatcatcatgataagcttgatgggaa	120		
Db 4365	tcagaagtttcacgacctacatctctcaagttatcatcatgataagcttgatgggaa	4424		
QY 121	gaagttgacacactatccaggaatccacactggaacctgaactatgctctctttggcaat	180		
Db 4425	gaagttgacacactatccaggaatccacactggaacctgaactatgctctctttggcaat	4484		
QY 181	ggattcatctctgggataaacaacatattttlaacctccaattatgtccgataactccg	240		
Db 4485	ggattcatctctgggataaacaacatattttlaacctccaattatgtccgataactccg	4544		
QY 241	tttgaccaccacacattatagcatctgcacgacatcttcgcatgagttgatgggcgtga	300		
Db 4545	tttgaccaccacacattatagcatctgcacgacatcttcgcatgagttgatgggcgtga	4604		
QY 301	tttaaatagtttcagcagcatgcccattgggaatgsgagtaagaacaatccagatgcacagat	360		
Db 4605	tttaaatagtttcagcagcatgcccattgggaatgsgagtaagaacaatccagatgcacagat	4664		
QY 361	tactgcttatccactattgaacaataatggttgcacactgtgtcccttcaaaagctcgact	420		
Db 4665	tactgcttatccactattgaacaataatggttgcacactgtgtcccttcaaaagctcgact	4724		

Db 6987 ccagagttggtgacacagattccctgagatgaggttctggtctgagagcacaga 7046
Qy 781 cctctactga 790
|||||
Db 7047 cctctactga 7056

RESULT 14
AAT31031
ID AAT31031 standard; cDNA; 8967 BP.
XX
AC AAT31031;
XX
DT 19-SEP-1996 (first entry)
XX
DE Factor-VIII full-length cDNA.
XX
KW Factor-VIII; retrovirus; vector; haemophilia A; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 110..7165
FT sig_peptide /*tag= a
FT sig_peptide 110..166
FT mat_peptide /*tag= b
FT mat_peptide 167..7162
FT /*tag= c
XX
PN MO621035-A2.
XX
PD 11-JUL-1996.
XX
PE 18-DEC-1995; 95MO-US16582.
XX
PR 30-DEC-1994; 94US-0366851.
XX
PA (CHIR) CHIRON VIAGENE INC.
XX
PI Bodner M, Chang S, Chi-Tang Hsu D, De Polo NJ;
XX
DR WPI: 1996-334010/33.
DR P-PSDB: AAM00465.
XX
PT Retroviral vector directing expression of full length factor VIII -
PT used in the gene therapy and treatment of haemophilia A
XX
PS Claim 3; Page 58-68; 86pp; English.
XX
CC A full-length cDNA clone (AAT31031) codes for human Factor-VIII
CC (AAM00465), a trace plasma glycoprotein which acts as a cofactor in
CC conjugation with Factor-IXa in the activation of Factor-X.
CC Retroviral vectors comprising the full-length cDNA can be efficiently
CC packaged into infectious retroviral particles. These may be used to
CC transduce cells either in vivo or ex vivo. Factor-VIII expressed
CC from such transduced cells will be processed and transported in a
CC fashion analogous to the expression product of a normal Factor-VIII
CC gene. Retroviral particles harbouring such vectors will be useful
CC in the gene therapy of haemophilia A.
XX
SQ Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;

Query Match 79.0%; Score 790; DB 17; Length 8967;
Best Local Similarity 99.4%; Pred. No. 3.6e-236;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 gatcaagtgatctgttggcacaatgatattacagcgatcaagaccaggtgccc 60
|||||
Db 6376 gatcaagtgatctgttggcacaatgatattacagcgatcaagaccaggtgccc 6435
|||||
Qy 61 tcagaagttccagccctacatctcagttatcatcatgatgtctgagtgaa 120
|||||

Db 6436 tcagaagttctccagccctacatctcagttatcatcatgatgtctgagtgaa 6495
Qy 121 gaatggcagacttatcgaggaattccactggaaccttaatggtctcttggcaatgt 180
|||||
Db 6496 gaatggcagacttatcgaggaattccactggaaccttaatggtctcttggcaatgt 6555
Qy 181 gattcatctgggataaacaacatatattttaacctccaattatgtctgataatccg 240
|||||
Db 6556 gattcatctgggataaacaacatatattttaacctccaattatgtctgataatccg 6615
Qy 241 ttgcaacccaactcattatagatctcgacgacctcttcgcatgagttgagtgctgta 300
|||||
Db 6616 ttgcaacccaactcattatagatctcgacgacctcttcgcatgagttgagtgctgta 6675
Qy 301 tttaaatgttcagcatgcatgttggaatggaatggaatgaatataatgaatgcaat 360
|||||
Db 6676 tttaaatgttcagcatgcatgttggaatggaatggaatgaatataatgaatgcaat 6735
Qy 361 tactgttcatctacttaccatattgtttgcaacctggtctccttcaaaagctcgact 420
|||||
Db 6736 tactgttcatctacttaccatattgtttgcaacctggtctccttcaaaagctcgact 6795
Qy 421 tcacctcaaggaggaagtaatgctctgagacctcagtgaaataatccaaagtgct 480
|||||
Db 6796 tcacctcaaggaggaagtaatgctctgagacctcagtgaaataatccaaagtgct 6855
Qy 481 gcaagtggacttccagaagaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 540
|||||
Db 6856 gcaagtggacttccagaagaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 6915
Qy 541 tctgcttaccagatgatgtgaagaggttctcctacatctcagcagtcagatgagcatca 600
|||||
Db 6916 tctgcttaccagatgatgtgaagaggttctcctacatctcagcagtcagatgagcatca 6975
Qy 601 gtgactctcttcttccagaatgcaagtaaggttttccagggaatccaagactcctt 660
|||||
Db 6976 gtgactctcttcttccagaatgcaagtaaggttttccagggaatccaagactcctt 7035
Qy 661 cacactgtggtgaactctctagaccaccggttactgactctactcctcgaattaccc 720
|||||
Db 7036 cacactgtggtgaactctctagaccaccggttactgactctactcctcgaattaccc 7095
Qy 721 ccagagttggtgacacagattccctgagatgagagttctggtctgagagcaagga 780
|||||
Db 7096 ccagagttggtgacacagattccctgagatgagagttctggtctgagagcaagga 7155
Qy 781 cctctactgactcagcc 798
|||||
Db 7156 cctctactgaggtgccc 7173

RESULT 15
AAC87526
ID AAC87526 standard; DNA; 8967 BP.
XX
AC AAC87526;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human factor VIII cDNA, SEQ ID NO:1.
XX
KW Factor VIII; human; A2 domain; C2 domain; LRP-mediated plasma clearance;
KW receptor-dependent clearance; receptor-independent clearance;
KW haemophilia; half-life; ss.
XX
OS Homo sapiens.
XX
PN MO200071714-A2.
XX
PD 30-NOV-2000.
XX
PF 24-MAY-2000; 2000MO-US14111.
XX

PR 24-MAY-1999: 99US-0135847.
 XX (AMN-) AMERICAN NAT RED CROSS.
 XX
 PI Saenko EL, Strickland DK;
 XX
 DR MPI: 2001-025163/03.
 DR P-SDB: AAB48843.
 XX
 PT Factor VIII mutants having increased half-life useful for treating
 PT hemophilia, comprise one or more amino acid substitutions in the A2
 PT and/or C2 domain of factor VIII.
 PS
 PS Claim 41: Page 86-101; 121pp; English.
 XX
 CC The invention relates to human factor VIII mutants comprising an amino
 CC acid substitution at one or more positions in the A2 domain and/or an
 CC amino acid substitution at one or more positions in the C2 domain.
 CC The invention also encompasses a factor VIII mutant which lacks a B
 CC domain (AAB48842). The factor VIII mutants have an increased half-life
 CC in the bloodstream. The A2 domain mutants exhibit reduced LRP-dependent
 CC (receptor-dependent) clearance of factor VIII, while C2 domain mutants
 CC have reduced receptor-independent clearance. The invention also relates
 CC to a method of using RAP (receptor associated protein), a protein which
 CC inhibits LRP (low density lipoprotein related protein)-mediated ligand
 CC internalisation, to increase the half-life of factor VIII. The mutant
 CC factor VIII proteins, and nucleotides encoding them, are useful
 CC for treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and
 CC nucleic acids encoding them may also be used in the treatment of
 CC haemophilia, in combination with a mutant factor VIII protein or DNA of
 CC the invention. The invention provides means of increasing the half-life
 CC of factor VIII by reducing its clearance from plasma. The present
 CC sequence represents cDNA encoding human factor VIII.
 XX
 SQ Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;

Query Match 79.0%; Score 790; DB 22; Length 8967;
 Best Local Similarity 99.4%; Pred. No. 3.6e-236;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgatctgttgacccaatgattatcagcgatcaagaccaggtgccg 60
 DB 6376 gatcaagtgatctgttgacccaatgattatcagcgatcaagaccaggtgccg 6435
 QY 61 tcagaagttccacgctctacatctcagttatcactgtatcattgtatgggaa 120
 DB 6436 tcagaagttccacgctctacatctcagttatcactgtatcattgtatgggaa 6495
 QY 121 gaagtgacagacttgcaggaattccactgaacttaagtctcttggcaatgt 180
 DB 6496 gaagtgacagacttgcaggaattccactgaacttaagtctcttggcaatgt 6555
 QY 181 ggaattcatctgggataaacaacatattttaaccctccaattattgtctgatacatcg 240
 DB 6556 ggaattcatctgggataaacaacatattttaaccctccaattattgtctgatacatcg 6615
 QY 241 ttgacccaacttaattagattcgagactcttcgatgagttgaggtctgtga 300
 DB 6616 ttgacccaacttaattagattcgagactcttcgatgagttgaggtctgtga 6675
 QY 301 tttaaatagttcagacatgcatctgggaatgagataaagaacatatacagatgcacagat 360
 DB 6676 tttaaatagttcagacatgcatctgggaatgagataaagaacatatacagatgcacagat 6735
 QY 361 taagttcatctactacttaccataatgttgcacactgtctctcttcaaaagctgact 420
 DB 6736 taagttcatctactacttaccataatgttgcacactgtctctcttcaaaagctgact 6795
 QY 421 tcacctcaaggagagataatgctgagagactcaggtgaataatccaaagatggtct 480
 DB 6796 tcacctcaaggagagataatgctgagagactcaggtgaataatccaaagatggtct 6855

QY 481 gcaatggaacttccagaagacatgaagtacagggagtactctcaggagataaatac 540
 DB 6856 gcaatggaacttccagaagacatgaagtacagggagtactctcaggagataaatac 6915
 QY 541 tctgcttaccagcatgtatgtgaaggagttcctatctccagcagtcacaatgagccatca 600
 DB 6916 tctgcttaccagcatgtatgtgaaggagttcctatctccagcagtcacaatgagccatca 6975
 QY 601 gtgaactctcttttcagaatgcaagtaaggtttttcagggaatacaagactcctt 660
 DB 6976 gtgaactctcttttcagaatgcaagtaaggtttttcagggaatacaagactcctt 7035
 QY 661 cacacctgtgtgaactctctagaacccagcttactactcgtactcttgaattacccc 720
 DB 7036 cacacctgtgtgaactctctagaacccagcttactactcgtactcttgaattacccc 7095
 QY 721 ccagagttggttgcacagattgcccctgagagatgtctctggtctgcagagcacagga 780
 DB 7096 ccagagttggttgcacagattgcccctgagagatgtctctggtctgcagagcacagga 7155
 QY 781 ccttactgactcagacc 798
 DB 7156 ccttactgaggtgtgccc 7173

Search completed: January 17, 2002, 16:34:34
 Job time: 12579 sec

1000

241 LLGcaccCcaCCLcaC

QY 241 ttgca

Query Match	Similarity	100.0%	Score 1000	DB 4	Length 4999
Best Local	Similarity	100.0%	Pred. No. 0		
Matches 1000	Conservative	0	Mismatches	0	Indels 0
					Gaps 0
QY	1	gacaaagtgatgacgtgttggcccaatgatattacgcgacatcaagcccaaggtgccc	60		
Db	4000	gacaaagtgatgacgtgttggcccaatgatattacgcgacatcaagcccaaggtgccc	4059		
QY	61	tcaagaagttctccagccctacatctcccaagttatcaatcatgatagctttagtgaa	120		
Db	4060	tcaagaagttctccagccctacatctcccaagttatcaatcatgatagctttagtgaa	4119		
QY	121	gaagtggcagacttatcgaggaattccacttggaaacttaatggtctctttggcaatgt	180		
Db	4120	gaagtggcagacttatcgaggaattccacttggaaacttaatggtctctttggcaatgt	4179		
QY	181	ggattcatctcgatataaacacaatatttttaacccctccaatttcttcgtatatccg	240		
Db	4180	ggattcatctcgatataaacacaatatttttaacccctccaatttcttcgtatatccg	4239		
QY	241	tttgacccaactcatattatagcatctcgacgacctcttcgcagtgaagttgagtgctga	300		

```

|||||
Db 4240 ttgaccccaactatcatagcctcgcagaccttcgcagatgagttgagtcgtga 4299
OY 301 tttaaatgctgacgacatgcatggaatgagatgaagaaacatgcagatcacagt 360
Db 4300 tttaaatgctgacgacatgcatggaatgagatgaagaaacatgcagatcacagt 4359
OY 361 tactgtctacatccttaacttaacataatgcttcgcacctggtctccttcaaaaagctgact 420
Db 4360 tactgtctacatccttaacttaacataatgcttcgcacctggtctccttcaaaaagctgact 4419
OY 421 tcaccttccaaaggagagatgatactcctgagaaactcaggtgataatccaaaagatgct 480
Db 4420 tcaccttccaaaggagagatgatactcctgagaaactcaggtgataatccaaaagatgct 4479
OY 481 gcaagtgtgactccaaagaacatgaagaatcacagagatgaactaccacaggaataaacc 540
Db 4480 gcaagtgtgactccaaagaacatgaagaatcacagagatgaactaccacaggaataaacc 4539
OY 541 tctgtctacacagatgatagtgaaaggatcctcactcactccacagatcgaagatgacatca 600
Db 4540 tctgtctacacagatgatagtgaaaggatcctcactcactccacagatcgaagatgacatca 4599
OY 601 gtgagctctctttttccagaatgcaaaagtattttccagggaatcaaacactcctt 660
Db 4600 gtgagctctctttttccagaatgcaaaagtattttccagggaatcaaacactcctt 4659
OY 661 cacacctgtgtgtaactctctctagacccacagctgactgactccttccttcgaatcaccc 720
Db 4660 cacacctgtgtgtaactctctctagacccacagctgactgactccttccttcgaatcaccc 4719
OY 721 ccagagttgggtgacacagattgacctgagatgagatgcttgggtcgtgaggaacagga 780
Db 4720 ccagagttgggtgacacagattgacctgagatgagatgcttgggtcgtgaggaacagga 4779
OY 781 cctctactgactgagccttaataaagaattattttcattgcaatagtgtgtgttt 840
Db 4780 cctctactgactgagccttaataaagaattattttcattgcaatagtgtgtgttt 4839
OY 841 ttgtgtgctggccgacaggaacccctgtagatgagttggccacctcctctctgcgcgcctc 900
Db 4840 ttgtgtgctggccgacaggaacccctgtagatgagttggccacctcctctctgcgcgcctc 4899
OY 901 gctcgtctacatgagcgcgagcgaccaaagttgcgccgaacgcccgggttgcgcgcgcg 960
Db 4900 gctcgtctacatgagcgcgagcgaccaaagttgcgccgaacgcccgggttgcgcgcgcg 4959
OY 961 cctcagttgagcgagcgagcgcgacgtccttcaggaat 1000
Db 4960 cctcagttgagcgagcgagcgcgacgtccttcaggaat 4999

```

RESULT 2

```

US-09-364-862-14
; Sequence 14, Application US/09364862
; Patent No. 6221349
; GENERAL INFORMATION:
; APPLICANT: Coulo, Linda B.
; APPLICANT: Colosi, Peter C.
; TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
; TITLE OF INVENTION: BY TARGET
; FILE REFERENCE: AVIGEN-03743
; CURRENT APPLICATION NUMBER: US/09/364,862
; EARLIER FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: 60/125,974
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/104,994
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 4999

```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-364-862-14

```

```

Query Match          100.0%; Score 1000; DB 4; Length 4999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 gatcaagtgagatctgttggcacaatgatatatcaacgcatcaagaacccagggtgcgcg 60
Db 4000 gatcaagtgagatctgttggcacaatgatatatcaacgcatcaagaacccagggtgcgcg 4059
OY 61 tcagaagttccacagccctcactatctcccaattatcatcatctgatatgtctgtggaa 120
Db 4060 tcagaagttccacagccctcactatctcccaattatcatcatctgatatgtctgtggaa 4119
OY 121 gaagtgagacattacgaggaattccacatggaaccttaattgtctcttcttgcaatgt 180
Db 4120 gaagtgagacattacgaggaattccacatggaaccttaattgtctcttcttggcaatgt 4179
OY 181 ggattcatctgggataaaacacaatatttttaacctccaattatgtctcgatacatcgcg 240
Db 4180 ggattcatctgggataaaacacaatatttttaacctccaattatgtctcgatacatcgcg 4239
OY 241 ttggacccaactcatatatacatttcgacagacctcttcgcagtgagttgagtgctgtga 300
Db 4240 ttggacccaactcatatatacatttcgacagacctcttcgcagtgagttgagtgctgtga 4299
OY 301 tttaaatgttgacagacatgacattggagatgagatgaagaaacatcacagatgcacgat 360
Db 4300 tttaaatgttgacagacatgacattggagatgagatgaagaaacatcacagatgcacgat 4359
OY 361 tactgtctacatccttaacttaccatatagtttgacacctggtctccttcaaaaagctgact 420
Db 4360 tactgtctacatccttaacttaccatatagtttgacacctggtctccttcaaaaagctgact 4419
OY 421 tcacctcaaaaggaggagtaatgctctgagagacccacagtgatataatccaaaagatgctc 480
Db 4420 tcacctcaaaaggaggagtaatgctctgagagacccacagtgatataatccaaaagatgctc 4479
OY 481 gcaagtgtgacttccagaagaacatgaaagtacacagaggttaactcctaaggagtaaatc 540
Db 4480 gcaagtgtgacttccagaagaacatgaaagtacacagaggttaactcctaaggagtaaatc 4539
OY 541 tctgtctacacagcatgatagtgaaaggatctcctacatctccagcagtcacaagtgcacatca 600
Db 4540 tctgtctacacagcatgatagtgaaaggatctcctacatctccagcagtcacaagtgcacatca 4599
OY 601 gtgagactcctttttcagaatgagcaagaatgattttcagggaataatcaagaactcctt 660
Db 4600 gtgagactcctttttcagaatgagcaagaatgattttcagggaataatcaagaactcctt 4659
OY 661 cacacctgtgtgtaactctctagacccacagctgactgactccttccttcgaatcaccc 720
Db 4660 cacacctgtgtgtaactctctagacccacagctgactgactccttccttcgaatcaccc 4719
OY 721 ccagaagtgggtgacacagattgcccctgagatggaagtctgtggtcgtgacgacagga 780
Db 4720 ccagaagtgggtgacacagattgcccctgagatggaagtctgtggtcgtgacgacagga 4779
OY 781 cctctactgactgagccttaataaagaattattttcattgcaatagtgtgtgttt 840
Db 4780 cctctactgactgagccttaataaagaattattttcattgcaatagtgtgtgttt 4839
OY 841 ttgtgtgctggccgacaggaacccctagtgatgagttggcactcctctcctgcgcgcgc 900
Db 4840 ttgtgtgctggccgacaggaacccctagtgatgagttggcactcctctcctgcgcgcgc 4899
OY 901 gctcgtcactgagcgcgagcgaccaaagttgcgccgaacgcccgggttgcgcgcgcg 960

```


Db 4900 gctgcgtactgagcgccgagccaaagctgcgccgagccggttcgcccggcg 4959
QY 961 cctcagtgagcgagcgagcgccagctgctgcagagcat 1000
Db 4960 cctcagtgagcgagcgagcgccagctgctgcagagcat 4999

RESULT 3

US-09-470-618-13
Sequence 13, Application US/09470618
Patent No. 6200560

GENERAL INFORMATION:

APPLICANT: Colosi, Linda B.
APPLICANT: Colosi, Peter C.

TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
TITLE OF INVENTION: by Target Cells

FILE REFERENCE: AVIGEN-04082

CURRENT APPLICATION NUMBER: US/09/470, 618

CURRENT FILING DATE: 1999-12-22

EARLIER APPLICATION NUMBER: 09/364, 862

EARLIER FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: 60/125, 974

EARLIER FILING DATE: 1999-03-24

EARLIER APPLICATION NUMBER: 60/104, 994

EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 2.0

SEO ID NO 13

LENGTH: 11933

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-470-618-13

Query Match 95.4%; Score 954.4; DB 4; Length 11933;
Best Local Similarity 97.9%; Pred. No. 4.3e-310;
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

QY 1 gatacaagtgatctgttggcaccatgattatcagcgacccaaggggtgccg 60
Db 3975 gatacaagtgatctgttggcaccatgattatcagcgacccaaggggtgccg 4034
QY 61 tcagaagttccagcctctacatctctcagttatcatatgatacttcttgaggaa 120
Db 4035 tcagaagttccagcctctacatctctcagttatcatatgatacttcttgaggaa 4094
QY 121 gaagtgagcagacttctcaggaattcaccatggaacttaagtgtctcttggaagt 180
Db 4095 gaagtgagcagacttctcaggaattcaccatggaacttaagtgtctcttggaagt 4154
QY 181 gattcatctggataaacaacataattttaaccctccaattatgtctcgatatactcg 240
Db 4155 gattcatctggataaacaacataattttaaccctccaattatgtctcgatatactcg 4214
QY 241 ttgcaaccaactcatatagatctgcagcactcttgcagatggatggaggtctgga 300
Db 4215 ttgcaaccaactcatatagatctgcagcactcttgcagatggatggaggtctgga 4274
QY 301 ttaataatgtcagacatgcattggaatggagataagaacatcatagatgcacaat 360
Db 4275 ttaataatgtcagacatgcattggaatggagataagaacatcatagatgcacaat 4334
QY 361 tactgtctcatctactatccaataatgttgcacactggtctcctcctcaaaagctcgat 420
Db 4335 tactgtctcatctactatccaataatgttgcacactggtctcctcctcaaaagctcgat 4394
QY 421 taacctccaaggagagatgctgtgagagactcaggtgtaataccaagaagtgcct 480
Db 4395 taacctccaaggagagatgctgtgagagactcaggtgtaataccaagaagtgcct 4454
QY 481 gcaagtgagctccaggaacaaatgaagtcacagagagtaactactcagggtgtaaatc 540

Db 4455 gcaagtgagctccaggaacaaatgaagtcacagagagtaactactcagggtgtaaatc 4514
QY 541 tctgtaccacagatgatagtgaagagttctcatalctccagcagatcaagatgacata 600
Db 4515 tctgtaccacagatgatagtgaagagttctcatalctccagcagatcaagatgacata 4574
QY 601 gtgactctcttttccaagaatggaaggttttcagggaatcaagactcctt 660
Db 4575 gtgactctcttttccaagaatggaaggttttcagggaatcaagactcctt 4634
QY 661 cacactgtgtgaactctctcagaccacacgcttaccagctcagctcctcgaattcacc 720
Db 4635 cacactgtgtgaactctctcagaccacacgcttaccagctcagctcctcgaattcacc 4694
QY 721 ccagaatgggtgacacagatgcccctgagatggaagtctgtgctgcagagcacaaga 780
Db 4695 ccagaatgggtgacacagatgcccctgagatggaagtctgtgctgcagagcacaaga 4754
QY 781 cctctactgactgagccttaataaggaaatttatctcagtaagtgtgtgtt 840
Db 4755 cctctactgactgagccttaataaggaaatttatctcagtaagtgtgtgtt 4809
QY 841 ttgtgtgcgcgcaggaacccctagtgatgagttggccactcctctctgcgcctc 900
Db 4810 ttgtgtgcgcgcaggaacccctagtgatgagttggccactcctctctgcgcctc 4869
QY 901 gctgcgtactgagcgccgagcacaaggtgcgcgcagccgagcttgcgcgcgcg 960
Db 4870 gctgcgtactgagcgccgagcacaaggtgcgcgcagccgagcttgcgcgcgcg 4929
QY 961 cctcagtgagcgagcgagcgccagctgctcctgagagat 1000
Db 4930 cctcagtgagcgagcgagcgccagctgctcctgagagat 4969

RESULT 4

US-09-364-862-13
Sequence 13, Application US/09364862
Patent No. 6221349

GENERAL INFORMATION:

APPLICANT: Colosi, Linda B.
APPLICANT: Colosi, Peter C.

TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
TITLE OF INVENTION: BY TARGET CELLS

FILE REFERENCE: AVIGEN-03743

CURRENT APPLICATION NUMBER: US/09/364, 862

CURRENT FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: 60/125, 974

EARLIER FILING DATE: 1999-03-24

EARLIER APPLICATION NUMBER: 60/104, 994

EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEO ID NO 13

LENGTH: 11933

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-364-862-13

Query Match 95.4%; Score 954.4; DB 4; Length 11933;
Best Local Similarity 97.9%; Pred. No. 4.3e-310;
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

QY 1 gatacaagtgatctgttggcaccatgattatcagcgacccaaggggtgccg 60
Db 3975 gatacaagtgatctgttggcaccatgattatcagcgacccaaggggtgccg 4034
QY 61 tcagaagttccagcctctacatctctcagttatcatatgatacttcttgaggaa 120

```

Db 4035 tcgaagttccagcctcactcctcagttatcatcatglatagcttgatggaa 4094
QY 121 gaatggcagactatcgaagaattccatctgaaccttaattgctttcttggcaatg 180
Db 4095 gaatggcagactatcgaagaattccatctgaaccttaattgctttcttggcaatg 4154
QY 181 ggaatcatctgggataaaacacataattttaaaccctcaattatgtctgatalaccg 240
Db 4155 ggaatcatctgggataaaacacataattttaaaccctcaattatgtctgatalaccg 4214
QY 241 ttggacccaactcatatagatctgcagcactcttcgcatgaggtgtaggctgtga 300
Db 4215 ttggacccaactcatatagatctgcagcactcttcgcatgaggtgtaggctgtga 4274
QY 301 tttaaatagttgagcatgcatctgggaatgaggtgaagaataatcagatgacagat 360
Db 4275 tttaaatagttgagcatgcatctgggaatgaggtgaagaataatcagatgacagat 4334
QY 361 tactgtctcactcaattatcaacataatgttgcacactgtctccttcaaaagctcag 420
Db 4335 tactgtctcactcaattatcaacataatgttgcacactgtctccttcaaaagctcag 4394
QY 421 tcacctccaaggaggagtaatgcttgcagcactcaggtgaatbaatccaaagatgagc 480
Db 4395 tcacctccaaggaggagtaatgcttgcagcactcaggtgaatbaatccaaagatgagc 4454
QY 481 gcaagtggaattccaaagaacaaatgaagaatcgaagagtaactcagaggtgtaaatc 540
Db 4455 gcaagtggaattccaaagaacaaatgaagaatcgaagagtaactcagaggtgtaaatc 4514
QY 541 tctgtctaccagcatgtagtgaagaaggttccatctcctcagcagtcgaatgagccatca 600
Db 4515 tctgtctaccagcatgtagtgaagaaggttccatctcctcagcagtcgaatgagccatca 4574
QY 601 gtagactccttttttcagaatgagcaaaagtttttcaaggaaatcaagactcctt 660
Db 4575 gtagactccttttttcagaatgagcaaaagtttttcaaggaaatcaagactcctt 4634
QY 661 caaacctgtgtgaactctcctcagacacacgttactgactcactccttcgaaatcaccc 720
Db 4635 caaacctgtgtgaactctcctcagacacacgttactgactcactccttcgaaatcaccc 4694
QY 721 ccaagatgtggtgacacagatttgccttgaggaatgaggttcttgagctgcgaagcaagga 780
Db 4695 ccaagatgtggtgacacagatttgccttgaggaatgaggttcttgagctgcgaagcaagga 4754
QY 781 ccttactgactcagacctaataaaggaaatttatttcaatgcaatagtggtgttt 840
Db 4755 ccttactgactcagacctaataaaggaaatttatttcaatgcaatagtggtgttt 4809
QY 841 ttgtgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 900
Db 4810 ttgtgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 4869
QY 901 gctgcgtcaactgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 960
Db 4870 gctgcgtcaactgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 4929
QY 961 cctcagttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1000
Db 4930 cctcagttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4969

```

RESULT

US-08-484-891-7

Sequence 7, Application US/08484891

Patent No. 5935935

GENERAL INFORMATION:

APPLICANT: Connolly, Sheila

APPLICANT: Kaleko, Michael

APPLICANT: Smith, Theodore

TITLE OF INVENTION: Adenoviral Vectors for

```

: TITLE OF INVENTION: Treatment of Hemophilia
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
: ADDRESSEE: Cecchi, Stewart & Olstein
: STREET: 6 Becker Farm Road
: CITY: Roseland
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07068
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch diskette
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: WordPerfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,891
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 514
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/218,335
: FILING DATE: 25-MAR-1994
: APPLICATION NUMBER: 08/074,920
: FILING DATE: 10-JUN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Olstein, Elliot M.
: REGISTRATION NUMBER: 24,025
: REFERENCE/DOCKET NUMBER: 271010-273
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-994-1700
: TELEFAX: 201-994-1744
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4629 bases
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna primer
: FEATURE:
: NAME/KEY: Factor VIII CDNA with
: NAME/KEY: B domain deleted
: US-08-484-891-7

```

Query Match 79.0%; Score 790; DB 2; Length 4629;

Best Local Similarity 99.4%; Pred. No. 4.4e-255;

Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 gataaaggtgagatcgttggacaaatgatatcagcgcatcaagaaccagggtgccg 60
Db 3627 GATCAAGGTGATCTGTGGACCAATGATTATTCACGGCATCAAGACCAGGTCGCCG 3686
QY 61 tcaagaatttcagcactcactcactcactcactcactcactcactcactcactcactcactc 120
Db 3687 TCAGAAAGTTCTCAGCCTCTACATCTCTCAGTTATCATCATATATGATGCTTATGCGAA 3746
QY 121 gaatggcagactatcgaagaattccatctgaaccttaattgctttcttggcaatg 180
Db 3747 GAAGTGCAGACTTATCGAAGAAATTCACCTGGAACCTTATATGCTTCTTGGCAATG 3806
QY 181 ggaatcatctgggataaaacacataattttaaaccctcaattatgtctgatalaccg 240
Db 3807 GATTTCATCTGGGATTAACAACAATATTTTAAACCTCCCAATTATGTCGATACATCGG 3866
QY 241 ttggacccaactcatatagatctgcagcactcttcgcatgaggtgtaggctgtga 300
Db 3867 TTGGACCCCAACTCATATTATGATTCGACACACTTTCGATGAGATTGATGGCTGTGA 3926
QY 301 tttaaatagttgagcatgcatctgggaatgaggtgaagaataatcagatgacagat 360
Db 3927 TTTAATAGTTGACGATGCGCATTTGGGAATGAGAAATTAAGCATATTCAGATGCACAGAT 3986
QY 361 tactgtctcactcaattatcaacataatgttgcacactgtctccttcaaaagctcag 420

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-717-294-41

Db 3987 TACTGCTTCATCTTACTTACCAATATGTTTGGCCACTGCTCTCTTCAAAAGTCTGACT 4046
Qy 421 tcacctcaaggaggaagtaatgctctgagacccaggtgaataatccaaaagatgct 480
Db 4047 TCACCTCCAAAGGAGAGTAATGCTTGAGACCTCAGGTGATATATCCAAAGAGTGGCT 4106
Qy 481 gcaagtgactccagaagaacatgaagatcacagagtaactctcaaggagtaaatc 540
Db 4107 GCAAGTGACTCTCCAGAAACATGAAGATGACAGAGTAACCTACTCAGGAGTAATATC 4166
Qy 541 tctgcttaaccgacatgatatgtaaggaggttcctcatctccagagtaaatgccaatca 600
Db 4167 TCTGCTTACCGACATGATGTGAAGAGTTCCTCATCTCCAGCATGAAGTGGCCATCA 4226
Qy 601 gtgactctctcttttcagaatggcaagtaaaagtttttcaggagaaatcaagactcct 660
Db 4227 GTGACCTCTCTTTTTCAGATGCGCAAAAGTAAAGTCTTTCAGGGAATCAAGCTCCTT 4286
Qy 661 caccactgtgtgaactctctagaaccacgcttaactgaactcgaacttgaatcaacc 720
Db 4287 CACACCTGTGGTGAACCTCTTACAGACCCACCTTACTGACTGCTACCTTCAAAATTCACCC 4346
Qy 721 ccagagttgggtgacccagatggccctgagagatgagagttctgggtcgagagcaagga 780
Db 4347 CCAGAGTTGGGTGACCGCATGATTGCCCTGAGATGAGAGTCTGGGGCTGCGAGGCACAGA 4406
Qy 781 cctctactgaactcgagcc 798
Db 4407 CCTTACTGAGGGTGGCC 4424

```

RESULT 6

US-08-717-294-41
 ; Sequence 41, Application US/08717294
 ; Patent No. 6114148

GENERAL INFORMATION:

APPLICANT: SEED, BRIAN

APPLICANT: HAAS, JURGEN

TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF

TITLE OF INVENTION: PROTEINS

NUMBER OF SEQUENCES: 110

CORRESPONDENCE ADDRESS:

ADDRESS: Clark & Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/717,294

FILING DATE: 20-SEP-1996

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/345001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX:

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 4670 base pairs

TYPE: nucleic acid

STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-717-294-41

Query Match 79.0%; Score 790; DB 3; Length 4670;
 Best Local Similarity 99.4%; Pred. No. 4.4e-255;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy 1 galcaagtgatctgttggcaacaaatgaattacagagcatcaagaccagatgccc 60
Db 3653 GATCAAGTGTGATCTGTGGACCAATGATTTATTCAGGCATCAAGACCAAGGTGCCG 3712
Qy 61 tcagaagttctccagccttaactctcaagttatcatcatgataagcttgatggga 120
Db 3713 TCAGAAGTTCTCCAGCCTCTACATCTCATGTTATCATCATGATATGATCTGATGGGA 3772
Qy 121 gaagtgacagactatcgaagaaatccactggaaccttaatgctctcttggcaatg 180
Db 3773 GAAGTGGACACTTATTCAGAGAAATTCACCTGACCTTAATGCTTCTTGGCAATGT 3832
Qy 181 ggaatcatctgggataaacaacataatttaaccctcaatattgtctcgataacccg 240
Db 3833 GGAATCATCTGGGATTAACACAAATATTTTAACCTCCAAATATTTGCTCGATACATCCG 3892
Qy 241 ttggaccccaactcattatagcatctgcagcaactcttcgcatggaatgagtgctga 300
Db 3893 TTGGACCCCAACTCATTTATATGATTCGACACTCTTCCATGGAATGATGAGGCTGTGA 3952
Qy 301 tttaaatagttgcagcatctgggaatggaagtaaaagcaatatcagatgcacagat 360
Db 3953 TTTAAATAGTTGCACGATGCGCATTTGGGAATGAGAGATGAAGCAATATCAGATCAGAT 4012
Qy 361 tactgtctacactacttaaccaatagtlttccacctgtgtcccttcaaaagctgaact 420
Db 4013 TACTGCTTCACTTACTTACCAATATGTTTGGCCACTGTCTCTTCAAAAGCTGAGCT 4072
Qy 421 tcacctcaaggaggaagatgctctggagacctcaggtgaataatccaaaagatgct 480
Db 4073 TCACCTCCAAAGGAGAGATGATGCTGAGACCTCAGGTGAATATTCAAAAGATGGCT 4132
Qy 481 gcaagtgactccagaagaacatgaagatcacaggaagtaactactcaaggagtaaatc 540
Db 4133 GCAAGTGACTTCCAGAGAACATGAAGATCACAGGAGTACTACAGGAGATAAATC 4192
Qy 541 tctgtctaccaagcatgtagtgaagaggttcctcaactccagcaagtcgaagtgcata 600
Db 4193 TCTGCTTACCGACATGATGTGAAGGATTCCTCATCTCCAGCATCAAGATGGCATCA 4252
Qy 601 gtgactctctttttcagaatgcaagtaaaagttttcaggagaaatcaagactcct 660
Db 4253 GTGACCTCTCTTTTTCAGATGCGCAAAAGTAAAGTCTTTCAGGGAATATCAACCTCTT 4312
Qy 661 caccactgtgtgaactcttagaacccacgcttactgaactcgtactcgtgaatcaacc 720
Db 4313 CACACCTGTGGTGAACCTCTTACAGACCCACCTTACTGCTGACTTGAATTCACCC 4372
Qy 721 ccagaagttgggtgacacagatgctcctgagatggaagttctggtctgagagcaagga 780
Db 4373 CCAGAAGTTGGGTGACACAGATGCTGCTGAGGATGAGAGTCTGGGCTGCGAGGCACAGA 4432
Qy 781 cctctactgaactcgagcc 798
Db 4433 CCTTACTGAGGGTGGCC 4450

```

RESULT 7

US-08-882-083-1

; Sequence 1, Application US/08882083

; Patent No. 5869292

; GENERAL INFORMATION:

; APPLICANT: VOORBERG, Johannes J.

TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,083
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5035 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 35..5017
US-08-882-083-1

Query Match 79.0%; Score 790; DB 2; Length 5035;
Best Local Similarity 99.4%; Pred. No. 4.6e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaaggtgagtcgttggtgacccaatgattatcaagggcatcagaccagggtgccc 60
DB 4231 GATCAAGGTGAGTCGTTGTTGGACCAATGATATTCACGGCATCAAGACCGAGGTGCCG 4290
QY 61 tcaagaattctccagcctctacatctctcaagttatcatcatgtaagctcttgatggaa 120
DB 4291 TCAGAAGTTCTCCAGGCTCTACATCTCTCAGTTTATCATCATGATGATGATGATGGA 4350
QY 121 gaagtgagcagacttatcgagggaatcaccctggaaccttaatgctctctcttggaatgt 180
DB 4351 GAAGTGGCAGACTTATCGAGGAATTCACCTGGAACCTTAAGGCTCTCTTGGGCAATGT 4410
QY 181 ggattcatctgggataaaacacaatatatttaacctccaattatctgctgatalacatcg 240
DB 4411 GGATTCACTTGCGATTAACACAAATATTTTAAACCTCCCAATTATGCTCGATACATCCG 4470
QY 241 ttgacccaactcatatagatctcgagcaactctcgcatgaggttgatgagctgtga 300
DB 4471 TTGACCCCACTCATATAGATCTCGACGACACTCTTCGCAATGAGTGTGATGGGCTGTGA 4530
QY 301 tttaaatgttgacagcatgcatcttggaatggagagtaagaacaatacagatgcaagat 360
DB 4531 TTTAAATAGTTTCACACATGCCATTGGGAATGAGATGAAGCAATATATCAGATGCACAGAT 4590
QY 361 tactgcttcactacttaacaatafatgttgccaccctggtctctcttaaaagctcgact 420
DB 4591 TACTGCTTCACTCTTACTTACCAATATATGTTGGCACCTGCTCTCTCTTAAAGCTCGACT 4650

QY 421 tcaactccaaggaaggaatgaatgccttgagagacctcgaagtgaaataatccaaaagatgact 480
DB 4651 TCACCTCCAAGGAGGAGATATGCTTGAGACCTCAGCTGAATATCCAAAGAGTGACT 4710
QY 481 gcaatggaacttccagaagaacatgaatgaatgcagaggaatgaactactcagggaatgaatc 540
DB 4711 GCAATGGACTTCCAGAGACAAATGAAGTCAACAGAGTAATCTCAGGAGATGAATATC 4770
QY 541 tctgcttaaccagcatgatatggaaggatctcattcaccagcaatgaatgagatca 600
DB 4771 TCTGCTTACCAAGCATGATGTGAAGGAGACTTCTCATCTCCAGCATCAAGATGCCATCA 4830
QY 601 gtgagactctcttttcagaatggcaagtaaaagtttttcaagggaataaagaactcct 660
DB 4831 GTGAGACTCTCTTTTTCAGAAATGCAAGATTAAGGTTTTCAGGGAATCAAGACTCCTT 4890
QY 661 cacacgttgatgaactctctagaccacacgcttactgactcgtactccttgaattacc 720
DB 4891 CACACTGTGTGAACCTCTCTAGACCCACGCTACTGACTGACTGACTGGAATTCACCC 4950
QY 721 ccagagttgggtgacacagatgcccctgagagatgaggttcttggtcgtcgaggaagga 780
DB 4951 CCAGAGTTGGGTGACACAGATTTGCCCTGAGAGATGAGGTTCTTGGGCTGCGAGGCACAGA 5010
QY 781 cctctactgactcgagcc 798
DB 5011 CCTCTACTGAGGTGGCC 5028

RESULT 8
US-08-558-107-1
Sequence 1, Application US/08558107
Patent No. 5910481
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,107
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5035 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 35..5017
US-08-558-107-1

Query Match 79.0%; Score 790; DB 2; Length 5035;
Best Local Similarity 99.4%; Pred. No. 4.6e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgatctgttgcacacatgattatccagggcaacaagccagggtccg 60
DB 4231 GATCAAGTGATCTGTGGCACCAGATATTATTCAGCGCATCAAGACCAGGGTCCCG 4230
QY 61 tcaagatctccagcctctacatctcagttatcatatgatatgttggaa 120
DB 4291 TCAGAACTTCCAGCCCTACATCTCTCAAGTTATCATATGATCTTGTATGGGA 4350
QY 121 gaagtggaagcttatgaggaatattcaactggaaccttaagtcttctggcaatgt 180
DB 4351 GAAGTGGAGACTTATTCAGGAATTCACAGTGAACCTTAAAGTCTTCTTGGCAATGT 4410
QY 181 gattcatctgggataaacaacatattttcaacctccaattatgtctgatacatccg 240
DB 4411 GATTCTATCTGGGATAAACAAATATTTTAAACCTCCATTTATGCTCATATCATCG 4470
QY 241 ttgacccaactcatatagcatctgcagacactcttcgcatgagtgatggcgtgta 300
DB 4471 TTTGACCCCACTCATATATAGATTCGACGACTCTTGCAATGAGTGTGAGTGGCTGGA 4530
QY 301 ttaaatagttgcagcatgcatctggagatggagatgaagaatataagatgacagat 360
DB 4531 TTTAAATAGTTGCAGCATGCCATGTTGGGAATGAGATAAGCAATATGATGCAAGAT 4590
QY 361 tactgtctcatctactcttaccacatgtttggccctgctctccccaagctcgact 420
DB 4591 TACTGCTTCATCTACTTATACCAATATGTTTGGCCACTGCTCTTCAAAAGCTCGACT 4650
QY 421 tcaactcagaaggagaaatgtaatgctgagacactcagtgaaataaccagaagtgct 480
DB 4651 TCACCTCCAGAGGAGAGATTAATGCTCGAGACCTCGAGTAAATATCCAAAGAGTGGCT 4910
QY 481 gaaatgtgactccaggaagacatgaagtaagtaactactcaggaagtaaatc 540
DB 4711 GCAAGTGGACTTCCAGAAACATGAATGTCACAGGAGTAACTACTCAGGGGTAAATTC 4770
QY 541 tctgtctacacagatgtaagagatcttccatctccagcaatcaagaatgagcaatca 600
DB 4771 TCTGCTTACACAGATGATGTAAGAGATCTCTCATCTCCAGCAATCAAGATGCGCATCA 4830
QY 601 gtgagactctcttttcagaaatgcaaaagttttcagggaaatcaagactcct 660
DB 4831 GTGAGACTCTCTTTTTCAGAAATGCAAGTAAAGTTTTCAGGGAATCAAGATCTCTCT 4890
QY 661 caaacctgtgtgaactctctagacacacacttactgactcgtctccttgaattaccc 720
DB 4891 CACACCTGTGTGAACCTCTAGACCCACCGTTACTGACTGCTTCAATTCGATTCACCC 4950
QY 721 ccagagttgggacacacagatgcccctagagatgaggtctcgggctcgaagacagaa 780
DB 4951 CCAGAGTTGGGTGCACACAGATGCTTGAAGATGAGAGTTCTGGGCTCGAGAGCAGAGA 5010
QY 781 cctctactgactcagacc 798
DB 5011 CCTCTACTGAGGTGGCC 5028

RESULT 9
US-09-243-539-1
; Sequence 1, Application US/09243539
; Patent No. 6130203
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,539
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5035 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 35..5017
US-09-243-539-1

Query Match 79.0%; Score 790; DB 3; Length 5035;
Best Local Similarity 99.4%; Pred. No. 4.6e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgatctgttgcacacatgattatccagggcaacaagccagggtccg 60
DB 4231 GATCAAGTGATCTGTGGCACCAGATATTATTCAGCGCATCAAGACCAGGGTCCCG 4290
QY 61 tcaagatctccagcctctacatctcagttatcatatgatatgttggaa 120
DB 4291 TCAGAACTTCCAGCCCTACATCTCTCAAGTTATCATATGATCTTGTATGGGA 4350
QY 121 gaagtggaagcttatgaggaatattcaactggaaccttaagtcttctggcaatgt 180
DB 4351 GAAGTGGAGACTTATTCAGGAATTCACAGTGAACCTTAAAGTCTTCTTGGCAATGT 4410
QY 181 gattcatctgggataaacaacatattttcaacctccaattatgtctgatacatccg 240
DB 4411 GATTCTATCTGGGATAAACAAATATTTTAAACCTCCATTTATGCTCATATCATCG 4470
QY 241 ttgacccaactcatatagcatctgcagacactcttcgcatgagtgatggcgtgta 300
DB 4471 TTTGACCCCACTCATATATAGATTCGACGACTCTTGCAATGAGTGTGAGTGGCTGGA 4530
QY 301 ttaaatagttgcagcatgcatctggagatggagatgaagaatataagatgacagat 360
DB 4531 TTTAAATAGTTGCAGCATGCCATGTTGGGAATGAGATAAGCAATATGATGCAAGAT 4590
QY 361 tactgtctcatctactcttaccacatgtttgcagacactcgtctcctcacaagctgact 420
DB 4591 TACTGCTTCATCTACTTATACCAATATGTTTGGCCACTGCTCTTCAAAAGCTCGACT 4650
QY 421 tcaactcagaaggagaaatgtaatgctgagacactcagtgaaataaccagaagtgct 480
DB 4651 TCACCTCCAGAGGAGATTAATGCTCGAGACCTCGAGTAAATATCCAAAGAGTGGCT 4910
QY 481 gaaatgtgactccaggaagacatgaagtaagtaactactcaggaagtaaatc 540

Db 4711 GCAAGTGGACTCCAGAGCAATGAAGTACAGAGAGTACTACTCAGGAGTAATAATC 4770
 QY 541 tctgtaccagcatgatgtgaaggaggtccatccatccacagcatgaatgacatca 600
 Db 4771 TCTGCTTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAGATGGCATCA 4830
 QY 601 gtgactctcttttttcaagaatgcaaaagtaaaagtttttcaaggaaatcaaatcctt 660
 Db 4831 GTGACACTCTCTTTTTCAGATAGGCAAAAGTAAAGTCTTTTCAGGAAATCAAGACTCTCT 4890
 QY 661 caacccgtgtgaactcttagaccacgcttactgactcgcttactcctgaattaccc 720
 Db 4891 CACACCTGTGTGAACCTCTAGACCCACCGTACTGACTGCTACTCTGCAATTCACCC 4950
 QY 721 ccagagttgggtgacacagattgcccctgaagatggaagttctgggctgcagagcacagga 780
 Db 4951 CCAGAGTTGGGTGACCAAGATTGCCCTGAGATGAGAGTTCTGGGCTGGAGGACACAGGA 5010
 QY 781 cctctactgactcgagcc 798
 Db 5011 CCTCTACTGAGGGTGCC 5028

RESULT 10
 ; US-08-276-594A-1
 ; Sequence 1, Application US/08276594A
 ; Patent No. 5693499
 ; GENERAL INFORMATION:
 ; APPLICANT: YONEMURA, Hiroshi
 ; APPLICANT: TAJIMA, Yoshitaka
 ; APPLICANT: SUGAMARA, Keishin
 ; APPLICANT: MASUDA, Kenichi
 ; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
 ; TITLE OF INVENTION: FACTOR VIII PROTEIN COMPLEX
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/276,594A
 ; FILING DATE: 18-JUL-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/950,191
 ; FILING DATE: 24-SEP-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 243262/1991
 ; FILING DATE: 24-SEP-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WEGNER, Harold C.
 ; REGISTRATION NUMBER: 25,258
 ; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6999 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS

LOCATION: 1..6996
 US-08-276-594A-1

Query Match 79.0%; Score 790; DB 1; Length 6999;
 Best Local Similarity 100.0%; Pred. No. 5,8e-255;
 Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaacaaggagatcgttggcacaatgatattcaagcgcatcaagaccagggtgccc 60
 Db 6210 GATCAAGGTGGATCTGTGGCCACCAATATTATTACGGCATCAAGACCCAGGGTGCCG 6269
 QY 61 tcaagaatctcagcctcctacatctcagtttcaatcatalcatalcatalcatalcatalc 120
 Db 6270 TCAGAAAGTCTCCACCGCTTACATCTCAGTTTATCATATGATATGATGAGGA 6329
 QY 121 gaagtgagcaacttctgaagaaatcacaacggaacctaagtcttcttggcaatgt 180
 Db 6330 GAAGTGGCAGACTTATCAGAGAAATCCACTGGAACCTTAATGGCTCTTGGCAATGT 6389
 QY 181 gattcaatcctgggataaacaacaataattttaaccctcaattatgtcgtatcatccg 240
 Db 6390 GATTCATCTGGGATAAACAACAATATTTTATACCTCTCAATTTATGCTCATATCATCCG 6449
 QY 241 ttggaccacaactcatalatagcatctcgagcactcttcgaatggatgaggtgagtg 300
 Db 6450 TTTGGACCCCAACTCATTTATAGCATTCGACAGACTCTGTCATGAGTGTGATGGGCTGTGA 6509
 QY 301 tttaaatgttgcagcagatgcatgtggaaatggagatgaagcaatatacagatgcacagat 360
 Db 6510 TTTAAATGTTGCACACATGCCATTGGGAATGAGAGTAAACCAATATCATGATGCACAGAT 6569
 QY 361 tactctcctacatcctccttaccatattgttgcacacgtggtcctccttaaaagtgcact 420
 Db 6570 TACTCTCTCATCTCTCTTACCAATATGTTTGCACACCTGCTCTCTTCAAAAGCTCAGCT 6629
 QY 421 tcaactcaaggaggaglaatgctcctggagacctcgaagtgaataatccaaaagatgact 480
 Db 6630 TCACCTCCAAAGGAGAGTAAATGCTCGAGACTCAGTGAATATCCAAAGAGTGGCT 6689
 QY 481 gcaagtgtgacttccagaagaatgaaagtcaacagaggaatcactcgaagagtaaatc 540
 Db 6690 GCAAGTGGACTTCCAGAGAACAAATGAAGTCAAGAGTAATCTCAAGGAGTAATAATC 6749
 QY 541 tctgttaccagcatgatgtgaaggagttcctcactccacagcagtaagaatgacatca 600
 Db 6750 TCTGCTTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCA 6809
 QY 601 gtgactctctttttcagaatggaagtaaaagtttttcaaggaaatcaagactcct 660
 Db 6810 GTGACACTCTCTTTTTCAGATAGGCAAAAGTAAAGTTCAGGGAATCAAGACTCTCT 6869
 QY 661 caacactgtgtgaactctcagaacccacgcttactgactcgcttactcctgaattaccc 720
 Db 6870 CACACCTGTGTGAACCTCTAGACCCACCGTACTGACTGACTGACTGACTGACTGACTGACT 6929
 QY 721 ccagagttgggtgacacagattgcccctgaagatggaagtttctgggctgcagagcacagga 780
 Db 6930 CCAGAGTTGGGTGACCAAGATTGCCCTGAGATGAGAGTTCTGGGCTGCCAGGACACAGGA 6989
 QY 781 cctctactga 790
 Db 6990 CCTCTACTGA 6999

RESULT 11
 ; US-08-121-202-1
 ; Sequence 1, Application US/08121202
 ; Patent No. 5563045
 ; GENERAL INFORMATION:
 ; APPLICANT: Pittman, Debra
 ; APPLICANT: Rehmetulla, Alnawaz
 ; APPLICANT: Wozney, John M.

APPLICANT: Kaufman, Randal J.
TITLE OF INVENTION: CHIMERIC PROCOAGULANT PROTEINS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
City: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,202
FILING DATE: 14-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Melner, M. C.
REGISTRATION NUMBER: 31,544
REFERENCE/DOCKET NUMBER: GI 5195A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 876-1210 X8574
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7056 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..7053
US-08-121-202-1

Query: Match 79.0%; Score 790; DB 1; Length 7056;
Best Local Similarity 100.0%; Pred. No. 5,8e-255;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcaagtgatctgttgacccaatgatattcaacggaatcaagaccaggtgccc 60
DB 6267 GATCAAGGTGATCTGTTGGACCAATGATTATTCACGGCATCAAGACCCAGGGTCCCG 6326
QY 61 tcagaagttccacccctcatctccatctatcatcatatgtagtcttgatgaa 120
DB 6327 TCAGAAGTTCCTCACCCCTCTCATCTCTCATATGATATGATATGATGATGAA 6386
QY 121 gaagtgacagacttcagaggaatctcacctggaaccttaatggtctctcttgcaatgt 180
DB 6387 GAAGTGCAGACTTTCAGAGGAATTCACAGGAACCTTAATGATGATCTTTGGCAATGT 6446
QY 181 ggattcatctgggataaacaacaatatatttaaccctcaattattctgcatatcatcg 240
DB 6447 GGATTCATCTGGGATMAAACAAATATTTTAAACCTTCCAATTTATGCTCATACATCCG 6506
QY 241 ttggacccaactcatatagcatctgcagacactctgcagatgagtgagtgctgtga 300
DB 6507 TTGGACCCCACTCTATTATGACTTCGACGACACTTTTGGCATGAGAGTTGATGGGCTGGA 6566
QY 301 tttaaatgttcagacatgcatcttggaatlgagaglaaagcaatatacagatgcacagat 360
DB 6567 TTTAATAGTTTCACAGATGCCATGTGGAGATGAGATGAAGCAATATACATGACAGAT 6626
QY 361 taactgtcatctcatcttaccataatgtttggcaactggtctctcttaaaagtgcagt 420
DB 6627 TACTGCTTCATCTTACTTACCAATATGTTGGCACCTGCTGCTCTTCAAAAGCTCGACT 6686

QY 421 tcacctccaagagagatgactcctgagaccctcagtgaaataatccaagaagtggct 480
DB 6687 TCACCTCCAAGGAGAGATGATGCTCGAGACCTCAGTGATTAATCCAAAAGAGTGGCT 6746
QY 481 gcaagtgaccttcagaaagaacaatgaaagtcacagaggaactaactcaaggagtaaatc 540
DB 6747 GCNAGTGCACTTCAGAAAGACATGAAGTCAACAGAGTACTACAGAGGAGTAAATC 6806
QY 541 tctgctaccagatgtagtgaagaggtctcctcatctccacagtcgaatggccatca 600
DB 6807 TCTGCTTACAGCATGATGATGAAGAGTTCTCATCTCCACAGTCAAGATGGCCATCA 6866
QY 601 gtgactctcttttcagaaatgcaagatggttttcagggaatcaagactcctt 660
DB 6867 GTGACTCTCTTTTTCAGAAATGCAAGTAAAGTTTTCAGGGAAATCAAGACTCCTT 6926
QY 661 caacactgtgtgaactctcttagaccaccgcttactgaactgcactcgaattcaacc 720
DB 6927 CACACCTGTGTAAGTCTCTAGACCCACCGTTACTGACTCGCTACCTTCGAATTCACC 6986
QY 721 ccagaattgggtgcacacagattgcccctgagatggaaggttctggtctgagagccagga 780
DB 6987 CCAGAGTTGGGTGCACACAGATTGCCCTGAGGATGAGGTTCTGGCTCGAGGACACAGSA 7046
QY 781 cctctactga 790
DB 7047 CCTTACTGA 7056

RESULT 12

US-08-366-851A-1
Sequence 1, Application US/08366851A

Patent No. 5681746

GENERAL INFORMATION:

APPLICANT: Bodner, Mordchai

APPLICANT: De Polo, Nicolas J.

APPLICANT: Hsu, David Chi-Tang

APPLICANT: Chang, Steven

TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Viagene, Inc.

STREET: 11055 Roselle Street

CITY: San Diego

STATE: California

COUNTRY: U.S.A.

ZIP: 92121

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/366,851A

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Chambers, Daniel M.

REGISTRATION NUMBER: 34,561

REFERENCE/DOCKET NUMBER: 930049, 438

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 452-1288

TELEFAX: (619) 452-2616

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 8967 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 110..7165

US-08-366-851A-1

Query Match 79.0%; Score 790; DB 1; Length 8967;
 Best Local Similarity 99.4%; Pred. No. 6.8e-255;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaagtgatctgttggcacaatgatattacacgcatcaagaaccaggtgccg 60
 Db 6376 GATCAAGTGGATCTGTGGCACAATGATTTTACGGCATCAAGACCAGGTCGCCG 6435
 QY 61 tcgaagtttccagcctcctcaactcctcagttatcatcatatgatagtcttgatgaa 120
 Db 6436 TCAGAAAGTTCCAGCCTTACATCTCTCAGTTATCATCATATGATTAAGCTTGATGGAA 6495
 QY 121 gaagttgcaactatcagagaatccacatggaactgaacttaagtctctcttgcaatgt 180
 Db 6496 GAAGTGGCAACTTATCGAGAAATTCACCTGACACTTATGCTTCTTTGGCAATGT 6555
 QY 181 ggaattcctgggataaacaacataatttaacccctcaatattgtctcgatataccg 240
 Db 6556 GGATTCATCTGGGATTAACAATATTTTAACTCCATTAATGTCGTGATACATCCG 6615
 QY 241 ttggcaccacacatattatagcattcgagcactcttcgcatgagttgagtgctgtga 300
 Db 6616 TTGCAACCCCAACTATTAATGATTCGACACCTTCCATGAGATGATGGCTGTGA 6675
 QY 301 tttaaatgattgagcagatgcatctgggaatggagatgaagaatcagatgagat 360
 Db 6676 TTTAATAGTTGACAGCATGCTGATGGAAATGAGATTAAGCAATATGCAATGACAGAT 6735
 QY 361 tactgtctacatcttacttaccataatgtttgacactgtctccttcaaaagctcgat 420
 Db 6736 TACTGCTTACTCTACTTATACCAATATGTTTGGCACCTGCTGCTCTTCAAAAGCTGACT 6795
 QY 421 tcacctccaaaggaggaatgaatgctcgagacactcagttgataatccaaagaatgctt 480
 Db 6796 TCACCTCCAAAGGAGAGATATGCTCGAGACCTCAGGTGAATTAATCCAAAGAGAGCT 6855
 QY 481 gcaagtggaactccagaagaacatgaagaatcagagagatgaactcagagagatgaatc 540
 Db 6856 GCAAGTGGACTTCCAAAGCAATGAAGTCAAGAGATGACTACTCAAGGAGTAAATC 6915
 QY 541 tctgtctacacgaatgatgatgaaggatcctcactcactcagcagatgagatgacatca 600
 Db 6916 TCTGCTTACAGCATGTATGTAAGAGTTCCATCTCCAGAGATGAGATGGCATCA 6975
 QY 601 gtggactctctttttcagaatgagcaagtaaaagttttcagggaatcaagaactcctt 660
 Db 6976 GTGGACTCTCTTTTTCAGAAATGGCAAAAGTTTTCAGGGAAATCAAGACTCCTT 7035
 QY 661 caaacctgtgtgaactctcctagaccaccgcttactgactgactcctcctggaatcacc 720
 Db 7036 CACACTGTGTGATCTCTTGAACCCAGCTTACTGACTGCTTCACTTCCGATTCACCC 7095
 QY 721 ccagagtttggtgacacagattgacctgagagatgagagttcctgggctgagagacagaa 780
 Db 7096 CCAGAGTTGGTGGACCAATATGCTCGAGATGAGAGTTCTGGGCTGGAGGACAGGA 7155
 QY 781 ccttactgactcagacc 798
 Db 7156 CCTTACTGAGGGGTGCC 7173

RESULT 13

US-07-864-004B-3

; Sequence 3, Application US/07864004B

; Patent No. 5364771

; GENERAL INFORMATION:

; APPLICANT: Lollar, John S.

; APPLICANT: Runge, Marshall S.

; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII

; NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kilpatrick & Cody

STREET: 1100 Peachtree Street

CITY: Atlanta

STATE: Georgia

COUNTRY: US

ZIP: 30309

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/864,004B

FILING DATE: 07 APRIL 1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Padst, Patrea L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: EMU106

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-815-6508

TELEFAX: 404-815-6555

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 9009 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Homo sapien

TISSUE TYPE: Liver

FEATURE:

NAME/KEY: misc.feature (Domain Structure)

LOCATION: 5001...7053

OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2"

US-07-864-004B-3

Query Match 79.0%; Score 790; DB 1; Length 9009;
 Best Local Similarity 99.4%; Pred. No. 6.8e-255;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaagtgatctgttggcacaatgatattacacgcatcaagaaccaggtgccg 60
 Db 6417 GATCAAGTGGATCTGTGGCACAATGATTTTACGGCATCAAGACCAGGTCGCCG 6476
 QY 61 tcgaagtttccagcctcctcaactcctcagttatcatcatatgatagtcttgatgaa 120
 Db 6477 TCAGAAAGTTCCAGCCTTACATCTCTCAGTTATCATCATATGATTAAGCTTGATGGAA 6536
 QY 121 gaagttgcaactatcagagaatccacatggaactgaacttaagtctctcttgcaatgt 180
 Db 6537 GAAGTGGCAACTTATCGAGAAATTCACCTGACACTTATGCTTCTTTGGCAATGT 6596
 QY 181 ggaattcctgggataaacaacataatttaacccctcaatattgtctcgatataccg 240
 Db 6597 GGATTCATCTGGGATTAACAATATTTTAACTCCATTAATGTCGTGATACATCCG 6656
 QY 241 ttggcaccacacatattatagcattcgagcactcttcgcatgagttgagtgctgtga 300
 Db 6657 TTGCAACCCCAACTATTAATGATTCGACACCTTCCATGAGATGATGGCTGTGA 6716

QY 301 tttaaatgttgcagcatgcattggaatggaagcaatcatcagatcacagat 360
Db 6717 TTTAAATAGTTCAGCATGCGATTGGAGATGAGATGAAGCAATATCAATGACAGAT 6776
QY 361 tactgttccacttaccatattgttgcacctgtctccttcaaaagctgact 420
Db 6777 TACTGCTTCACTTCTTACCAATATGTTGCCACCTGGTCTCTTCAAAAGCTGACT 6836
QY 421 tcacctcaagagagatlaatgcctgagagctcaggtgaataatccaaagagtgct 480
Db 6837 TCACCTCCAAAGGAGATATGCTGGAGACTCAGGTGAATATCCAAAGAGTGCT 6896
QY 481 gcaagtgtacttccagaagaacatgaaagtacagagagtaactctcagaggttaaatc 540
Db 6897 GCAAGTGACTTCCAGAAAGCAATGAAGTCAAGAGTAATCTCAGGAGTAATATC 6956
QY 541 tctgttccagagatgtaagtgaagaggttctcctcctcagagcaagtgagca 600
Db 6957 TCTGCTTACAGCATGTATGGAAGAGATTCTCATCTCCAGCATCAAGTGCCATCA 7016
QY 601 gtgacctcttlttccagaatggcaagtaaggttttccaggaaatcaagctcctt 660
Db 7017 GTGAGCTCTCTTTTTCAGATGCAAGTAAGTTCAGGGAATCAAGCTCCTT 7076
QY 661 cacacctgtgtgaactctctagaaccacgttactgactcgtcacttgaattcaacc 720
Db 7077 CACACCTGTGTGAATCTCTTGAACCCAGCTTACTGATCTGATCTTCAATTCACCC 7136
QY 721 ccaggtgtggtcagagatctgcctgaagatgaggttctggtcgtcagagacagga 780
Db 7137 CCAGGTGTGTTGGTGCACCATGTTGCCCTGAGATGAGGTCTGGGCTGCCAGCAGGA 7196
QY 781 cctctactgactgagcc 798
Db 7197 CCTTACTGAGGAGTGCC 7214

RESULT 14
US-08-251-937A-3
Sequence 3, Application US/08251937A
Patent No. 5583209
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marshall S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/251,937A
FILING DATE: 31-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EMU106DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6367
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 9009 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: liver
FEATURE:
NAME/KEY: misc_feature (Domain Structure)
LOCATION: 5001..7053
OTHER INFORMATION: /note="Equivalent to the A3-C1-C2
OTHER INFORMATION: domain"
FEATURE:
NAME/KEY: misc_feature (Domain Structure)
LOCATION: 1..2277
OTHER INFORMATION: /note="Equivalent to the A1-A2
OTHER INFORMATION: domain"
US-08-251-937A-3

Query Match 79.0%; Score 790; DB 1; Length 9009;
Best Local Similarity 99.4%; Pred. No. 6.8e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgatctgttgcacaaatgattatccagcatcaagcaagaggtgcgg 60
Db 6417 GATCAAGTGTGATCTGTGGCACCATTGATTATTCAGCATCAAGCAAGAGTGC 6476
QY 61 tcagaagtttccagacctcaatctcagttatcatcatgataagcttgatggaa 120
Db 6477 TCAGAAGTCTCCAGCCTTCAATCTCTCAGTTATCATCATATGATGAGGAA 6536
QY 121 gaagtgacagactatcagaagaattccactggaaccttaagtctcttggcaatg 180
Db 6537 GAAGTGCAGACTTATCGAAGAAATTCACCTGGAACCTTATGCTTCTTTGGCAATG 6596
QY 181 ggaattcctgggaataaacaacataatttaacctccaattatgtcgtatataccg 240
Db 6597 GGAATCATCTGGGATTAACAACATATTTTAACCTCCCAATATGTCGATACATCCG 6656
QY 241 ttgacccaactcatatagcatctgcagcaactcttcgatatgagttgagtcgtga 300
Db 6657 TTTGCACCCCAACTATATAGCATTCGACACTCTTCCGATGAGTGTGATGGCTGTGA 6716
QY 301 tttaaatgttgcagcatgcattggaatggagagaaagcaataatcagatcacagat 360
Db 6717 TTTAAATAGTTCAGCATGCGATTGGAGATGAGATGAAGCAATATCAATGACAGAT 6776
QY 361 tactgttccacttaccatattgttgcacctgtctccttcaaaagctgact 420
Db 6777 TACTGCTTCACTTCTTACCAATATGTTGCCACCTGGTCTCTTCAAAAGCTGACT 6836
QY 421 tcacctcaagagagatlaatgcctgagagctcaggtgaataatccaaagagtgct 480
Db 6837 TCACCTCCAAAGGAGATATGCTGGAGACTCAGGTGAATATCCAAAGAGTGCT 6896
QY 481 gcaagtgtacttccagaagaacatgaaagtacagagagtaactctcagaggttaaatc 540
Db 6897 GCAAGTGACTTCCAGAAAGCAATGAAGTCAAGAGTAATCTCAGGAGTAATATC 6956
QY 541 tctgttccagagatgtaagtgaagaggttctcctcctcagagcaagtgagca 600
Db 6957 TCTGCTTACAGCATGTATGGAAGAGATTCTCATCTCCAGCATCAAGTGCCATCA 7016
QY 601 gtgacctcttlttccagaatggcaagtaaggttttccaggaaatcaagctcctt 660
Db 7017 GTGAGCTCTCTTTTTCAGATGCAAGTAAGTTCAGGGAATCAAGCTCCTT 7076

Qy	661	caaccctggtgtaactctctagaccacccgttaactgactgctactcgtcaattacc	720
Db	7077	CACACCTGTGGTGAACCTCTTAACACCCACCGTTACTGACTCCCTCACTTTCGAAATTCACCC	7136
Qy	721	ccagagttgagtgaccacagatgcccctgagatgagtgactgagctgcgcagacacaga	780
Db	7137		7196
Db	7137	CCAGAGTTGGGTGCACCAATGATCCCTGAGGATGAGGATTTCTGGGCTCGAGGACACAGGA	7196
Qy	781	ccctactactgactgagacc	798
Db	7197		7214
		CCCTCTACTGAGGCTGGCC	7214
		RESULT 15	
		US-08-212-133A-1	
		/ Sequence 1, Application US/08212133A	
		/ Patent No. 5663060	
		/ GENERAL INFORMATION:	
		/ APPLICANT: Lollar, John S.	
		/ APPLICANT: Runge, Marshall S.	
		/ TITLE OF INVENTION: Hybrid Human/Animal Factor VIII	
		/ NUMBER OF SEQUENCES: 12	
		/ CORRESPONDENCE ADDRESS:	
		/ ADDRESSEE: Kilpatrick & Cody	
		/ STREET: 100 Peachtree Street	
		/ CITY: Atlanta	
		/ STATE: Georgia	
		/ COUNTRY: US	
		/ ZIP: 30303	
		/ COMPUTER READABLE FORM:	
		/ MEDIUM TYPE: Floppy disk	
		/ COMPUTER: IBM PC compatible	
		/ OPERATING SYSTEM: PC-DOS/MS-DOS	
		/ SOFTWARE: Patent In Release #1.0, Version #1.25	
		/ CURRENT APPLICATION DATA:	
		/ APPLICATION NUMBER: US/08/212,133A	
		/ FILING DATE: March 11, 1994	
		/ CLASSIFICATION: 435	
		/ PRIOR APPLICATION DATA:	
		/ APPLICATION NUMBER: US 07/864,004	
		/ FILING DATE: 07-APR-1992	
		/ ATTORNEY/AGENT INFORMATION:	
		/ NAME: Pabst, Patrea L.	
		/ REGISTRATION NUMBER: 31,284	
		/ REFERENCE/DOCKET NUMBER: EMU/76677	
		/ TELECOMMUNICATION INFORMATION:	
		/ TELEPHONE: 404-572-6508	
		/ TELEFAX: 404-572-6555	
		/ INFORMATION FOR SEQ ID NO: 1:	
		/ SEQUENCE CHARACTERISTICS:	
		/ LENGTH: 9009 base pairs	
		/ TYPE: nucleic acid	
		/ STRANDEDNESS: single	
		/ TOPOLOGY: linear	
		/ MOLECULE TYPE: cDNA	
		/ HYPOTHEICAL: NO	
		/ ANTI-SENSE: NO	
		/ ORIGINAL SOURCE:	
		/ ORGANISM: Homo sapien	
		/ TISSUE TYPE: liver	
		/ FEATURE:	
		/ NAME/KEY: misc_feature (Domain Structure)	
		/ LOCATION: 5125 . . . 7053	
		/ OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2	
		/ OTHER INFORMATION: domain"	
		/ FEATURE:	
		/ NAME/KEY: misc_feature (Domain Structure)	
		/ LOCATION: 1 . . . 2277	
		/ OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain."	
		/ FEATURE:	
		/ NAME/KEY: Domain	
		/ LOCATION: 1..2277	
		/ OTHER INFORMATION: /note= "cDNA encoding human factor	

OTHER INFORMATION: VIII.
US-08-212-133A-1

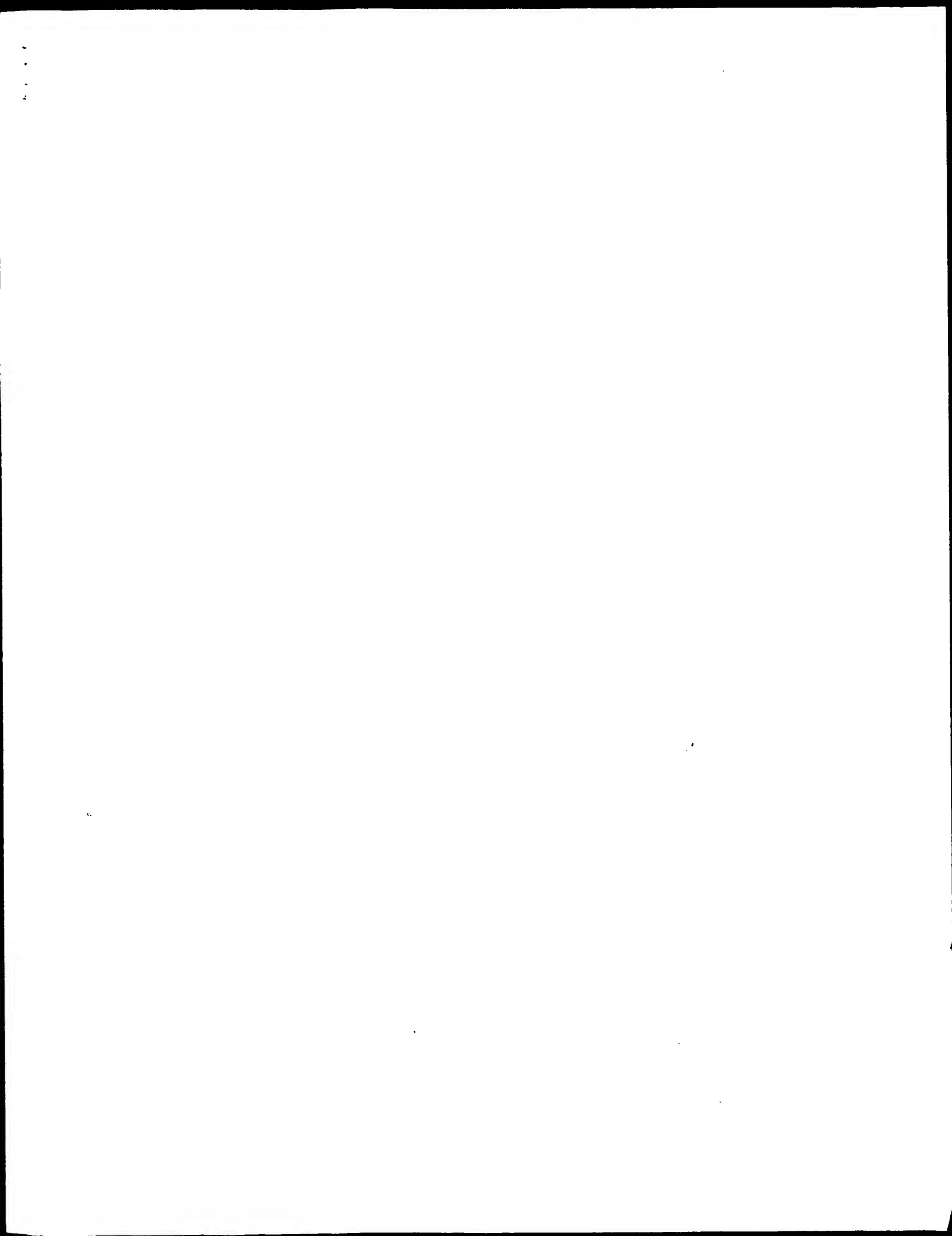
Query Match	79.0%;	Score 790;	DB 1;	Length 9009;
Best Local Similarity	99.4%;	Pred. No. 6.8e-255;		
Matches 793;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;

QY	1	gaanaagtgagctcgtttggcaccaaatggattatcaagcgcatcaagaccagagtgccg	60
Db	6417	GATTAAGGTGATCTGTGGCACCAATGATTTATTCAGGCAACCAAGCCAGGTGCCG	6476
QY	61	tcaagaattctccagcctctacatctctcagttatcatcatcgtatagcttgagaa	120
Db	6477	TCAGAAATTCCAGCCTTACATCTCTCAGTTTATTCATCATGTATAGTCTGATGGAA	6536
QY	121	gaagatgcagacttatcagaggaattccacctggaaccttaatggtctcttgcaatgt	180
Db	6537	GAACTGGCAACATTATCGAGAAATTCACCTGGAACTTTAATGGTCTTTTGGCAAGT	6596
QY	181	ggatctatcgggataaaacacatatltttaaaccctcaattatgtctcgataaccg	240
Db	6597	GGATTCTCTGGGATTAACACAAATTTTTTAACCTCCAAATTAATGTCTCATATATCCG	6656
QY	241	tttgcaccaactcatatagcatctcgacacactcttcgaltgagttgagtgcgtga	300
Db	6657	TTTGCACCCCAACTCATATTAATGATGTGCGACACTCTTCCGATGGAGTTGATGGCTGGA	6716
QY	301	tttaaatagttgcagcatgcacattggaattggaatgaagataatacatatgcacagat	360
Db	6717	TTTTAAATAGTTGGAGATGGCAATTTGGGAATGAGATGAAGTAATGCATAATGCACAAAT	6776
QY	361	tactgcttactactacttaccacatatgtttgcacactggtctccttcaaaagctcgact	420
Db	6777	TACTGCTTACTTACTTATACCAATATGTGTGCCACACTGTCTCCTTCAAAAGCTCACT	6836
QY	421	tcaacttcaaggaggagataatgctctgagacctcaagtgaataatccaagaagtgtct	480
Db	6837	TCACCTCCAAAGGAGAGATTAATGCCCTGGACCTCAGTGATTAATCCAAAAGAGTGGCT	6896
QY	481	gcaagtgagacttccagaagaacaaatgaagaatcaccagagatgaactactcaggaagttaaatc	540
Db	6897	GCAGATGGAATTCCAAAGACATGAATAATGCACAGAGTAATCTACTCGGGAGTAAATTC	6956
QY	541	tctgtcttaccagcatgtatgtgaaggagttccctcatctccagcagtcacaagtgtgcacatca	600
Db	6957	TCGTCTTACAGCATGTATGTGAAGAGATTTCTCATCTCCACAGTCAAGATGGGCATCA	7016
QY	601	gttgagactctcttttccagaatgycaaagttaaaagtttttcaaggaaatcaagactcctt	660
Db	7017	GTTGAGACTCTTTTTCACAATATGGCAAAATAAAGTTTTTTCAGGAAATCAAGACTCCTT	7076
QY	661	caaacctgtgttaactctctagaccacacgcttaactgactcgtaaccttgaattcaacc	720
Db	7077	CACACCTGTGTAACTCTCTAGAACCCACCGTTACTGACTGTGCTACCTTGAATTCACCC	7136
QY	721	ccagaagtgggtgcacagaattgccttgagatgagttctgtgctgcgcagacaga	780
Db	7137	CCAGAGTGTGGTGCACCAATATGCCCTGAGATGAGAGGTTCTGGGCTCCGAGCACAGA	7196
QY	781	cctctactgactgcagcc 798	
Db	7197	CCTCTACTGAGGCTGCC 7214	

Search completed: January 17, 2002, 16:28:08
Job time: 12398 sec

Tue Jan 22 15:12:47 2002

us-09-740-211-14_copy_4000_4999.rni



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 14:58:55 ; Search time 3755.78 Seconds
(without alignments)
2861.131 Million cell updates/sec

Title: US-09-740-211-14_COPY_4000_4999

Perfect score: 1000
Sequence: 1 gatcaagctgcatctgtgtg.....cgcagctgctgcagagacat 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthm:*
3: em_estlin:*
4: em_estom:*
5: em_estopl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_estl:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vtl:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length DB	ID	Description
1	573.8	57.4	692	11	BG700655 602682272
2	506	50.6	720	10	AL601022 DKFZP313G
3	449.2	44.9	815	10	AU135340 AU135340
4	446.4	44.6	943	11	BG818325 602780212
5	398.8	39.9	482	10	BE650938 BE650938
6	350.8	35.1	515	10	AV748467 AV748467
7	280.4	28.0	639	11	BI067808 BI067808
8	217.4	21.7	401	11	BE847128 BE847128
9	201.4	20.1	444	10	AW489638 UT-M-BH3-
10	194	19.4	444	10	AA184901 AA184901
11	193.4	19.3	384	11	BE852200 BE852200
12	193.2	19.3	247	10	AW46518 AW46518

c	13	190	19.0	431	10	AW456831	UT-M-BH3-
c	14	158.8	15.9	927	10	AL552613	AL552613
c	15	136.6	13.7	745	10	AL577036	AL577036
c	16	136.2	13.6	965	11	BP302800	BP302800
c	17	132.6	13.3	616	11	BG641736	BG641736
c	18	131.8	13.2	864	11	AL324553	AL324553
c	19	130.6	13.1	942	11	BI173222	BI173222
c	20	130	13.0	657	10	AL173222	AL173222
c	21	129	12.9	1038	11	BE911604	BE911604
c	22	128	12.8	772	11	BI251160	BI251160
c	23	127.8	12.8	290	11	BE84985	BE84985
c	24	127	12.7	857	11	BG242396	BG242396
c	25	125.6	12.6	775	11	BG246010	BG246010
c	26	125.4	12.5	608	11	BE554133	BE554133
c	27	121.8	12.2	961	11	BI249890	BI249890
c	28	121.4	12.1	555	11	BG086480	BG086480
c	29	120.6	12.1	674	11	BI409154	BI409154
c	30	120.6	12.1	613	10	AA512671	AA512671
c	31	119.8	12.0	811	11	BI154893	BI154893
c	32	119.6	12.0	698	10	AL575850	AL575850
c	33	118.8	11.9	763	11	BG176281	BG176281
c	34	117.4	11.7	633	10	AA208846	AA208846
c	35	116.4	11.6	984	11	BG176124	BG176124
c	36	115.8	11.6	654	11	BE916565	BE916565
c	37	115.8	11.6	916	11	BG242699	BG242699
c	38	115.2	11.5	616	10	BE284409	BE284409
c	39	114.4	11.4	270	11	BG090843	BG090843
c	40	114.2	11.4	784	11	BI105565	BI105565
c	41	112.4	11.2	794	11	BI154670	BI154670
c	42	112.4	11.2	871	11	BF167957	BF167957
c	43	111.4	11.1	790	11	BG976390	BG976390
c	44	110.4	11.0	549	10	BE381295	BE381295
c	45	110.4	11.0	626	11	BI065410	BI065410

ALIGNMENTS

RESULT 1
LOCUS BG700655 692 bp mRNA
DEFINITION 602682272F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4815208 5',
mRNA sequence.
ACCESSION BG700655
VERSION BG700655.1 GI:13970214
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10711 row: 0 column: 17
High quality sequence stop: 690.
Location/Qualifiers
1..692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4815208"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"

FEATURES

Source

```

/lab.host="DH10B"
/Note="Organ: Brain; Vector: pBluescriptPK (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagc
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to R0r 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NIHGR1, National
Institutes of Health). Note: this is a NIH-MGC Library."
162 a 165 c 173 g 172 t

```

Query Match	57.48	Score 573.8	DB 11	Length 692
Best Local Similarity	99.38	Pred. No. 2.3e-158		
Matches 597	Conservative 0	Mismatches 2	Indels 2	Gaps 2

Oy 161 atgctcctccttggcaatgttgattcatctcggataaaacacaatatlttaacctcca 220
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 91 AAGGCTTCTTTGGCAATGTGATTTCATCTGGATAAACAACACATATTTTAACCTCCA 150

Oy 221 atattgcctgataacatccgtttgcaccccaactcaattagcatcgcgaaccttgcg 280
 |||||
Db 151 ATATTGCTCGATACATCCGTTTGCAACCACAATCATTAAGCATTGGCACACTTTCGC 210

QY 281 atcgagttcaatggtctgtgattaaagtgcagcatgcatggaatgagagtaaa 340
|||||
Db 211 atgcagttgatggcctgatttaattagttgcagcagccattgggaatggagactaaa 270
|||||

OY	341	gcacatacagatgcaacagtactgtctcatccacttaaccaatatgttggcaccttg	400
Db	271	GCAATATCAGATGCACAGATTACTGCTTCATCCCTACTTACCATAATGTTTGCCACTGG	330

OY	401	tctccttcaaaagctcgacttcacccccaaggaggaatbctctggaacctcagtgtg	460
Db	331	TTCCTTCAAAAGCTCGACTTCACCCCTCCAAGGAGGAATATGCCTGGACACCTCAGGTG	390

QY 461 aataatccaaaagagtgycgtcgaagtgacttcagaagacaaatgaagtcacagagta 520
|||||
Db 391 AATATCCAAAGAGTGGCTGCAGTGGACTTCGAGAACAATGAAAGTCACAGAGTA 450

Qy 521 acctactcagygagtaaatctctgctaccagcatgtatgtga-ggagttcccatctc 579
|||||
Db 451 acctactcagggagtaaatctctgcttaccagcatgtatgtgaacggaggtcccatctc 510

QY 580 cagcagtcagaatgycatcagtgagctctctttttcagaatgycaaagta-aagttt 638
|||||
Db 511 CACcAGTCAGATGCCCATcAGTGCCTCTCTTTTTCAGAAATGCCAAATCAAGTCT 570
|||||

QY 639 ttcagggaatcaagactccttcaacccctgtgtgaactctctagaccaccggttactga 698
|||||
Db 571 ttcagggaatcaagactccttcaacccctgtgtgaactctctagaccaccggttactga 630

QY 699 ctgcgtacctgcgaattcacccccagatctggtggtgcaccagatctgccttgagatgag 758
|||||
db 631 ctgcgtacctgcgaattcacccccagatctggtggtgcaccagatctgccttgagatgag 690

QY	759	t	759
Pb	691	T	691

RESULT 2
A1601022

LOCUS	720 bp	EST	14-AUG-2001
DEFINITION	DKFZP313G1439_r1 313 (synonym: h1cc2) Homo sapiens cDNA clone		
DESCRIPTION	DKFZP313G1439 5', mRNA sequence.		
ACCESSION	AF011022		

```

VERSION      AL601022.1  GI:15164528
KEYWORDS     EST.
SOURCE       human.

```

TITLE	JOURNAL	REFERENCE	AUTHORS
ESR (Duesterhoeft, et al.) Unpublished (1999) Contact: Duesterhoeft A		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 720) Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J., and Wiemann, S.	

MIPS
Am Klopferplatz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the Human Genome Project.

This clone (DKFZp313G1439) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```
FEATURES      Location/qualifiers
source        1. .720
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp313G439"
/clone_1lb="313 (synonym: hlcc2)"
/dev_stage="adult"
/lab_host="DH10B"
/notes="vector: pTR1plex2; site_1: SfilA; site_2: SfilB"
cdna_collection="
178 a 186 c 166 g 189 t 1 others

```

Query Match	50.6%	Score 506;	DB 10;	Length 720;
Best Local Similarity	99.8%;	Pred. No. 2.4e-18;		
Matches 506;	Conservative	0;	Mismatches 1;	Indels 0;
			Gaps	0;

Oy 284 gagttagatggtctgtgatttcaatagttgcagcatgcatgtggaaatgagagttaaagca 343
 |||||
 Db 2 GAGTTGATGGGCTGTGATTAAATAGTTGCAGCATGCCATTGGGAAATGGAGAGTAAAGCA 61

QY	344	atatacagatgcacagattactcgtctcatctacttaccataatgtttgccacctggtct	403
Db	62	ATATCAGATGCACAGATTACTCGTCTTACCTCTTTACCAATATGTTTGGCACCCTGGTCT	121

QY 404 ccttcaaaagctcgacttcactccaaggagagtaatgctgtgagacctcagtgat 4633

Db 122 CCTTCAAAAGCTCGACTTCACTCCAAGGAGAGTAATGCTGTGAGACCTCAGTGAAT 1811

QY 464 aatccaaaagagtgcctgcgaagtgcacttcacagaacatgaatgaatccacgaagtact 5233
|||
Db 182 aatccaaaagagtgcctgcgaagtgcacttcacagaacatgaatgaatccacgaagtact 2411
|||

QY 524 actcaaggagtaaaatctctcttaccagcatgtatgtgaaggagttcctcatctccagc 583

QY 584 agtcaagatgycacatcagtgcactctcttttttcagaatgycaaagttaaaggttttcag 643
|||||
db 302 AGCTCAAGAGATGCGCCATTCAGTGTACACTCCTCTTTTGGTGAGAAATGCCAAGAATAAGCTGTTGTCAG 361

644 ggaatcaagactcccttcacacctgtgtgaactctctagaccacgcgttactgactgc 703

QY 704 taccttgcgaattcaaccccaagatttggtgtgcaaccagattgccctgagatgaggtctctg 763

QY 764 ggctgcgagcacagacctctactga 790
|||||

RESULT 3
LOCUS AU135340 815 bp mRNA EST 24-OCT-2000
DEFINITION AU135340 PLACE1 Homo sapiens cDNA clone PLACE1001786 5', mRNA
sequence.
ACCESSION AU135340
VERSION AU135340.1 GI:10995879
KEYWORDS EST.
SOURCE human.
ORANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 815)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Ishigai, T.
AUTHORS HRI human cDNA project
TITLE Unpublished (2000)
JOURNAL
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Source Location/Qualifiers
1..815
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1001786"
/clone_lib="PLACE1"
/tissue_type="placenta"
/note="Vector: pWE18SFR3"
BASE COUNT 200 a 207 c 204 g 197 t 7 others
ORIGIN
Query Match 44.9%; Score 449.2; DB 10; Length 815;
Best Local Similarity 98.3%; Pred. No. 1.5e-121;
Matches 454; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 161 atggtctctcttggaatgattcatctggtatgaataacacacatattttaaccctcca 220
DB 92 AAGGCTCTTCTTGCGAATGATGATTCATCTGGATAAACAAATATTTTAACCTTCCA 151
OY 221 attattgctcgatataccgtttgacccaactcatatagacattcgacacattctgcg 280
DB 152 ATTATGTCGATACATCCGTTTGACCCAACTCATATATTCATTCGACGACATCTTGC 211
OY 281 atggaggttgatggctggtgatattaatgattgacacatgacattggtggaatgagataa 340
DB 212 ATGGAGTTGATGGCTGTGATTTAAATGATTGACACATCCATTGGATGGAAATGAGATGAA 271
OY 341 gcaatacagatgacacattactgcttcatctacttaccacatagtttgacactg 400
DB 272 GCAATATGATGATGACAGATTAATCTCTTACATCTTACCAATATGTTTGCCACTGG 331
OY 401 tctcttcaaaagctcgacattcaactccaaggagagatgattgctggaagccttaagt 460
DB 332 TCTCTTCAAAAGCTCGATTCACCTCCAAAGGAGAGATTAATGCTTGAGACCTCAGGTG 391
OY 461 aataatcaaaagatgctgcaagtgcattccagaagacacatgaagtcacagagata 520
DB 392 AATAATCCAAAAGAGTGGCTGCAATGGAATCCCAAGAACATGAAGTCAACAGAGATA 451
OY 521 actactcagagataaactctctgcttaccagacatgatatgtaagagattctcattcc 580
DB 452 ACTACTCAGGAGTAAATCTCTGCTTACACAGCATGTATGTGAAGAGTCTCTCATCTCC 511

OY 581 agcagtcagatggtccatcagtggaactctcttttccagaat 622
DB 512 AGCAGTCAGATGCGCATCATGATGACTCGCTTACTTTCGAATT 553
RESULT 4
LOCUS BG818325 943 bp mRNA EST 22-MAY-2001
DEFINITION BG818325 602780212F1 NCL_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4915634
5', mRNA sequence.
ACCESSION BG818325
VERSION BG818325.1 GI:14165912
KEYWORDS EST.
SOURCE human.
ORANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 943)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
AUTHORS Contact: Robert Strausberg, Ph.D.
JOURNAL Email: cgabbs-r@mail.nih.gov
COMMENT Tissue Procurement: David N. Louis, M.D.
cDNA library preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLNL0824 row: h column: 03
High quality sequence stop: 629.
FEATURES
Source Location/Qualifiers
1..943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4915634"
/clone_lib="NCL_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV-SPO6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dfr.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP library."
BASE COUNT 252 a 241 c 210 g 240 t
ORIGIN
Query Match 44.6%; Score 446.4; DB 11; Length 943;
Best Local Similarity 99.8%; Pred. No. 1.1e-120;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 343 aatlcagatgacacattactgcttcatctacttaccacatagtttgacactggtc 402
DB 2 AATATGATGATGACAGATTAATCTGCTTACATCTTACCAATATGTTTGCCACTGTC 61
OY 403 tctcttcaaaagctcgacattcaactccaaggagagatgattgctggaagccttaagt 462
DB 62 TCTCTTCAAAAGCTCGATTCACCTCCAAAGGAGAGATTAATGCTTGAGACCTCAGGTGAA 121
OY 463 taatcaaaagatgctgcaagtgcattccagaagacacatgaagtcacagagataac 522
DB 122 TAATCCAAAAGAGTGGCTGCAATGGAATCCCAAGAACATGAAGTCAACAGAGATA 181
OY 523 tactcagagataaactctctgcttaccagacatgatatgtaagagattctcattccag 582
DB 182 TACTCAGGAGATAAATCTCTGCTTACACAGCATGTATGTGAAGAGTCTCTCATCTCCAG 241
OY 583 cagtcagagatgacatcagtggaactctcttttccagaatgagacagtaagtttttca 642
DB 242 CAGTCAGAGATGCGCATCATGATGACTCTCTTTTTCAGAAATGGCAAAAGTAAAGTTTTC 301

QY 643 gggaaatcaagactctctcaccacgtgtgtgaactctctagaccacccgttactgctc 702
 |||||||
 Db 302 GGGAAATCAAGACTCTCTTCACTGTGCGTGAAGTCTCTAGACCCACCGTACTGACTCG 361
 |||||||
 QY 703 ctactctgaattccaccacccaggttggtgacccagattgcccctgaagatgagttct 762
 |||||||
 Db 362 CTACCTCTGCAATTCACCCCGAGTTGGGTGACACAGATTGCCCCGTGAGTGGAGGTTCT 421
 |||||||
 QY 763 gggctgcagagcagacagactctactga 790
 |||||||
 Db 422 GGGCTGCGAGGACACAGACTCTACTGA 449
 |||||||
 RESULT 5
 BE650938 482 bp mRNA EST 06-SEP-2000
 LOCUS BE650938
 DEFINITION UI-M-BH3-asm-g-04-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
 UI-M-BH3-asm-g-04-0-UI 5', mRNA sequence.
 ACCESSION BE650938
 VERSION BE650938.1 GI:9976762
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 482)
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov
 cDNA library preparation: M. B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESEARCH GENERICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements
 Seq primer: M13 Reverse.
 FEATURES
 Source
 1. 482
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-asm-g-04-0-UI"
 /clone_1id="NIH_BMAP_M_S4"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_M_S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH_BMAP_M_S4,
 NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
 NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
 (NIH_BMAP_M_S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
 NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1

libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH_BMAP_M_S4 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996).
 BASE COUNT 142 a 102 c 101 g 137 t
 ORIGIN
 Query Match 39.9%, Score 398.8; DB 10; Length 482;
 Best Local Similarity 89.2%; Pred. No. 9.8e-107;
 Matches 430; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 QY 113 gatggagaagatggcagacttaccgagaatccacatgaacttaagtctcttc 172
 |||||||
 Db 1 GATGGGAGAGAGTGGCTGAGTATCCAGGAATTCACCTGGAACCTTAATGTTTCTTT 60
 |||||||
 QY 173 ggcattgtgattcatctgggataaacaataatttttaaccccaattatgcga 232
 |||||||
 Db 61 GGCATTGTGACTCATCTGGGATTAAGCATATAGTTTAACTCCCAATTATGCTCGA 120
 |||||||
 QY 233 tacatcgttcgaccacacattatagcatctcgacacactctcgatgagttgatg 292
 |||||||
 Db 121 TATATCGCTTGGACCCACCATCTTACGATCCGTAAGTCTTCCGATGAGATTATG 180
 |||||||
 QY 293 ggcctgtattaaatagttgacatgcatctgggaatlgagaglaaagcaatcatcat 352
 |||||||
 Db 181 GGCCTGTGATTTAAACAGTTCGACATACCATTTGGGAATGGAATGAATGAATATGAT 240
 |||||||
 QY 353 gcaagattatgcttactctacttaccataatgtttgcacactgctctctcaaaa 412
 |||||||
 Db 241 ACACAAATCAGCTGCTCTCTACTCTCCACCAACATGTTTGTACTTGTGCTCTTCACAA 300
 |||||||
 QY 413 gctgacttcacccccaagagaggaatgctccctggagacctcagtgtaataatccaaa 472
 |||||||
 Db 301 GCTGACTTTCACCTCCAGGAGGAGCATATGCTCTGGGAGCTCAGTGGAATGATCCAAA 360
 |||||||
 QY 473 gagtgtcgaagtggacttcagagaacaaatgaagaaatcagaagtaactacagga 532
 |||||||
 Db 361 CAATGGTTGCAAGTGAGCTTCAAAAGACATGAAGTCACTGGAAATTAATACCCAGGA 420
 |||||||
 QY 533 gtaaatctcgtctcagcatatgtagaagatgctctcattcccaagctcaaat 592
 |||||||
 Db 421 GTGAATCTCTTTACCGACATGTTGTGAAGAGTTCTTATTTTCACGACGACAGAT 480
 |||||||
 QY 593 gg 594
 ||
 Db 481 GG 482
 ||
 RESULT 6
 AV748467 515 bp mRNA EST 19-OCT-2000
 LOCUS AV748467
 DEFINITION AV748467 NPC Homo sapiens cDNA clone NPCXAA05 5', mRNA sequence.
 AV748467
 ACCESSION AV748467
 VERSION AV748467.1 GI:10906315
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 515)
 REFERENCE
 1 Song, H., Peng, Y., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N.,
 Qian, B., Liu, F., Ou, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu
 W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Ye, M., Zhang, Q.,
 Han, Z., Chen, Z., Hu, R. and Chen, J.
 Homo sapiens NPC library cDNA clones
 Unpublished (2000)
 TITLE
 JOURNAL
 COMMENT
 Contact: Qinghua Zhang
 Shanghai Institute of Endocrinology, Rui-Jin Hospital

197 Rui-Jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64370045(ex.663332)
Fax: 86-21-64743206
Email: mshl@ms.sh.cn

This clone is available at Shanghai Hematology Institute in
Shanghai.
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.

FEATURES

source

Location/Qualifiers
1. 515

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1b="NPCXAA05"
/clone_1b="NPC"
/tissue_type="pituitary"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 132 a 114 c 131 g 135 t 3 others
ORIGIN

Query Match 35.1%; Score 350.8; DB 10; Length 515;
Best Local Similarity 95.3%; Pred. No. 1.5e-92;
Matches 383; Conservative 0; Mismatches 14; Indels 5; Gaps 2;

OY 161 atgctctcttgcgaatgctgattcattcctggtgataaacaacattttaaaccctcca 220
DB 111 AAGGTCTCTTTGGCAATGATGATTCATCTGGGATAAACAATATTTTAAACCTCCCA 170
OY 221 attatgctgatacatcgctttgacccaactcattatgacattgcagacattctgc 280
DB 171 ATATCTCTCGATACACCTGTTGACACCAACATATATGACATTCGACACCTCTCGC 230
OY 281 atggaagtgtgagctgtgtttaaagtgtgacgacattggaggaagagataaa 340
DB 231 ATGAGATTGATGGGCTGTGATTTAAATGTTGACGATGCGCATTTGGAAATGAGAGATAA 290
OY 341 gcaatatacatgacagatctactctcattcacttaccataatglttgcacctg 400
DB 291 GCATATTCAGATGACAGATTAAGTTCATTCCTTACCAATATGTTGGCACCTCG 350
OY 401 tctcttcaaaagctgactcactcctcaaggagagagatgctgtgaaactcagtg 460
DB 351 TCTCCTTCAAAAGCTGACCTTCAAGGAGAGATGCTGGAACCTCAGAGTG 410
OY 461 aataatcaaaagagtgctgacgactgtcagagaagaacatgaatgcacagagata 520
DB 411 AATATATCCAAAGAGTGGCTTCAAGTGCAGCTTCAGAAACATGNGTCACA-NGAAGTA 469
OY 521 actactcaaggagataaactctgcttaccagacatgattgtg 562
DB 470 ACTACTTAAGGAGTA---CTCTCTTACCAAGCATGTATGTG 507

RESULT 7

BI067808

LOCUS

DEFINITION

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

pflin.pk008.h22

REFERENCE 1 (bases 1 to 639)

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

ADDRESS

CITY

STATE

COUNTRY

UNIVERSITY

DEPARTMENT

TELEPHONE

FAX

EMAIL

WWW

FTP

GATEWAY

MAILING

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

DISCUSSION

LISTSERV

RESULT	8			
LOCUS	BE847128			
DEFINITION	BE847128	401 bp	mRNA	EST
ACCESSION	U62211.1			26-SEP-2000
VERSION	5			
KEYWORDS	'similar to gb:L05573 Mus domesticus coagulation factor VII mRNA', complete (MOUSE);, mRNA sequence.			
SOURCE	BE847128			
ORGANISM	EST.	GI:10305395		
	house mouse.			
	Mus musculus			
	Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 401)			
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov This clone is available royalty-free through LLNL; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. MG:1093217 Seq primer: -40BP from GIBCO.			

	FEATURES
	source
	Location/Qualifiers
	1..401
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/clone_image="3417405"
	/clone_id="Soares mouse 3NBMS"
	/sex="male"
	/tissue_type="Spleen"
	/dev_stage="4 weeks"
	/lab_host="DH10B"
	/note="vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dn) primer [5' TGTTACCAATGCTGAAGTCGGAGCGCCGCCGTCTTTTTTTTTTTTTTTT 3'] : double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Benito Soares and M.Fatima Bonaldo."
BASE COUNT	117 a
ORIGIN	95 c 77 g 112 t

Query Match	21.7%	Score 217.4	DB 11	Length 401
Best Local Similarity	84.2%	Pred. No. 3.5e-53		
Matches 245	Conservative	0	Mismatches 46	Indels 0
				Gaps 0
QY	500	acatgaagaatcacagagtaactactcaaggaggaatcaatctcgtctaccagcatgat	559	
DB	2	ACATGGAATCATCTGTGATTAATAACCCAGGAGTGAATTTCTTTTACACGATGTTT	61	
QY	560	gtgaagagttctcactctccagcagatcaagatgycatcagttgactcctttttcag	619	
DB	62	GTCGAAGAGTCTCTTAATTTCCAGCAGTCACAGATGGCCATCATCTGATCAAAATTTTATAC	121	
QY	620	aatgccaagaatgaagtttttctaggagaaatcaagaactcttcacactgtgtgaactc	679	
DB	122	AATGGCAAGGTAAAGTCTTTTTCAGGGGATACAGAGCTCATCAACACTATGATGAATTC	181	
QY	680	cttagaccaccgcttaactgaactcgccttaacttgaattcaacccccagaagttgggtgacag	739	
DB	182	CTAGACCCACCAATTAATCATCTGCTATTTTGAATTCAACCCCCAGATCTGGGAGCACCA	241	
QY	740	atgagccctgaagtagaggtcttgagtcgcgcgagcaacgaagcctcctatga	790	
DB	242	ATTGCTCTGAGGCTTGAGATTCTAGAGATGTAGGCCACGACGAATATCTTA	292	

RESULT	9			
AM489638/c				
LOCUS				
DEFINITION	AM489638	444 bp	mRNA	EST
	UI-M-BH3-asm-g-04-0-UI.s1	NH_BMAP_M.S4	Mus musculus	cDNA clone
	UI-M-BH3-asm-g-04-0-UI.3,			mRNA sequence.
ACCESSION	AM489638			
VERSION	AM489638.1	GI:7059908		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	I (bases 1 to 444)			
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.			
TITLE	Normalization and subtraction: two approaches to facilitate gene			
	discovery			
JOURNAL	Genome Res.	6 (9),	791-806	(1996)
MEDLINE	97044477			
COMMENT	Contact: Chin, H			

National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag served to identify the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized brain stems library cDNA library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMP CDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMP CDNAs whose availability will be considered under appropriate and limited collaborative arrangements. The following repetitive elements were found in this cDNA sequence: 18-106, >CORR1#ULR/MALR
Seq primer: M13 Forward
POLYA-tes.

FEATURES

source

1. 444

Location/Qualifiers

1. 444

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="U1-M-BH3-asm-g-04-0-U1"

/clone_11b="NIH_BMAP_M_S4"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker. Site 1: Not I. Site 2: Eco RI. The NIH_BMAP_M_S4 library is a subtracted library of a series ultimately derived from a mixture of individually tagged (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S4, NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the

NIH BMAP M.S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG_LIB=NIH_BMAP_M_S4
TAG_TISSUE=brain-stems
TAG_SEQ=TCATG"

BASE COUNT 122 a 80 c 100 g 142 t
ORIGIN

Query Match 20.1%; Score 201.4; DB 10; Length 444;
Best Local Similarity 82.5%; Pred. No. 1.9e-48;
Matches 245; Conservative 0; Mismatches 46; Indels 6; Gaps 1;

494 cagaagacaatgaatgcagaggaagtaactactcaaggagtaaaatctctgtaccagc 553
DB 444 CAAAAGACAAAGAAAGTCATCGAATATAATACCAAGGAGAGAAATCTCTTTACAGC 385
554 atgtatgtgaaggaagttctcatctccagcagtcagaatgagtcagtgactcttt 613
DB 384 ATGTTTGGAAAGAGCTTCTTATTTCCAGCAGTCAAGATGCCCTCATGACTCAATTT 325
614 ttccagaatggcaagtaaggttttcaaggaaatcaagactccttcacactgtgtg 673
DB 324 TTATACAAATGCG-----AAGGTTTTCAGGGGATGACGACATCCACACCTATGATG 271
674 aactcttagaccacagcttactcgtactcgtactcgtactcgtactcgtactcgt 733
DB 270 AATTCTTAGACCCACCACTTACTCATCTGCTATCTTCAATTCACCCCAATGCGAG 211
734 caccagatgcctcagagatgaggtcttggtcgcagagagcagagactactga 790
DB 210 CACCAATTCCTCTGAGGCTTGAGATTTAGAGATGTAGGCCCAAGCATACTGA 154

RESULT 10

AA184901 265 bp mRNA EST 17-FEB-1997
LOCUS
DEFINITION mE3e12.r1 Soares mouse 3nBMS Mus musculus cDNA clone IMAGE:67486
5' similar to gb:L05373 Mus domesticus coagulation factor VIII mRNA
, complete (MOUSE);, mRNA sequence.

ACCESSION
VERSION AA184901
KEYWORDS
SOURCE AA184901.1 GI:1768759
EST.

ORGANISM
house mouse.
Mus musculus

REFERENCE
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucada, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE The Mashu-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:389478

Seg primer: -28M13 rev2 from Amersham
High quality sequence from: 250.

FEATURES

1. .265
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"

/clone="IMAGE:637486"
/clone.lib="Soares mouse 3nBMS"
/sex="male"
/tissue_type="Spleen"
/tissue="4 weeks"
/dev_stage="4 weeks"
/lab_host="D10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer (5',
TGTTCACATCTGAAAGTGGAGCGCCGCTGTTTCTTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."

BASE COUNT 82 a 63 c 53 g 67 t
ORIGIN

Query Match 19.4%; Score 194; DB 10; Length 265;
Best Local Similarity 86.3%; Pred. No. 2.5e-46;
Matches 226; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

375 acttaccatattgttgcacactgtctccttcaaaagctcactcactccaagga 434
DB 1 ACTTCACCAATGTTTGTCTTACTTGTGCTCTTCCACAGTGCATTCACCTCAGGAA 60
435 ggaatgctcgtgagacactcagtgtaataatccaaaagagtgctgcaagtgtctc 494
DB 61 GGACTAATGCTGCGCCACCCACAGTGAATATCCAAACAAATGTTGCAAGT-GACTTAC 119
495 agaaagaaatgaatgacagaggaagtaactactcagagaggaatctcgtttacaga 554
DB 120 AAAAGACAAATGAAGTCTCTGGAATATAACCAAGGAGGAATCTCTTTACAGCA 179
555 tgtatgtgaaggaagttcctcatctccagcagtcagaatgagtcactgtactcttt 614
DB 180 TGTTTTGGAAAGTCTTCTTATTTCCAGCAGTCAAGATGCCATCTGACATCAATTT 239
615 ttcagaatggcaagtaaggt 636
DB 240 TATACATGCGCAAGGTAAAGT 261

RESULT 11

BE852200/c 384 bp mRNA EST 26-SEP-2000
LOCUS
DEFINITION uw22b11.x1 Soares mouse 3nBMS Mus musculus cDNA clone IMAGE:3417405
3' similar to SW:PA8_MOUSE Q06194 COAGULATION FACTOR VIII PRECURSOR
, mRNA sequence.

ACCESSION
VERSION BE852200
KEYWORDS BE852200.1 GI:10310464
EST.

ORGANISM
house mouse.
Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 384)

TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT other ESTs: uw22b11.y1
Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1093217.

FEATURES

1. .384
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"

```

/db_xref="taxon:10090"
/clone="IMAGE:3417405"
/clone.lib="Soares mouse 3NDKs"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
/Note="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15',
5'-TGTACCAATCTGAAGTGGAGCGGCCGCGCTTTTCTTTTCTTTTCTTTT
3'; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT773 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT      110 a      73 c      91 g      110 t
ORIGIN

```

```

Query Match      19.3%; Score 193.4; DB 11; Length 384;
Best Local Similarity 84.2%; Pred. No. 4.2e-46;
Matches 218; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```

```

QY 532 agtaaaatctctgtaccagcatgtatgtgaaggagttccatctccagcagtaaga 591
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 384 AGGAATATCTCTTTACGACGATGTTGGAAGATTCCTATTTCACGACGATCAAA 325
QY 592 tggccatcagtgactctcttttccaagaatgcaagaagttttccaaggaaata 651
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 324 TGCCATCATCTGACCTCAATTTTATACATGCGCAAGTAAAGTTTTCAGGGGATCA 265
QY 652 agactctctacacactgtgtgaactctctagaccacccgttactcactcgtcactctg 711
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 264 GGACTATCATCACTATGATGATTTCTAGACCCACCATTAATCTAGCTATCTTGG 205
QY 712 aattccccccagagttggtgacagatttgccttgagatgagttcttggtctgca 771
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 204 AATTACACCCCAATGCGGAGCACAAAATGCTGTGAGCTTGAAGATTCGAGATGA 145
QY 772 ggcacagagactctactga 790
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 144 GGCCACGACGACATCTGA 126

```

RESULT 12

```

AM446518      247 bp      mRNA      EST      25-APR-2001
LOCUS      86227 MARC IBOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION      AM446518
ACCESSION      AM446518
VERSION      AM446518.1 GI:6988305
KEYWORDS
SOURCE
ORGANISM      COV.
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 247)

```

```

REFERENCE
AUTHORS      Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Caass,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
TITLE      Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL      Genome Res. 11 (4), 626-630 (2001)
MEDLINE      21180013
COMMENT
CONTACT: Smith TPJ
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

```

```

Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACATCAGCAGC
Plate: 48 row: F column: 20
Seq primer: ATTAGGACACTATG.
FEATURES
SOURCE
1..247
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone.lib="MARC IBOV"
/tissue_type="pooled"
/lab_host="DH10B"
/Note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
BASE COUNT      56 a      68 c      61 g      62 t
ORIGIN

```

```

Query Match      19.3%; Score 193.2; DB 10; Length 247;
Best Local Similarity 86.6%; Pred. No. 4.2e-46;
Matches 213; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```

```

QY 520 aactctcagagagtaaatctctgtctaccagcatgtatgtgaaggagttcccatc 579
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 AACTACCCCAAGGGGAATCTGCTTACGACATGATGTGAAGAGTTCTCATATC 61
QY 580 cagcagtaaatgagccatcagtgactctcttttccaagaatgcaagaagtttt 639
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 CAGTATGTAAGATGCGCATTAATCTGACATCTTCTTCAAGATGCAAAAGTGAAGTTT 121
QY 640 tcaaggaaatacagaactctcttcaacactgtgtgaactctctagaccacccgttactgac 699
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 122 TCAGGGAATCAAGACTCTCTACCCCGTGTGAATGCTCTAGACCCCGCTGTTAC 181
QY 700 tgcctactctgaattcaacccccagagttggtgacagattgcccctgagatgaggt 759
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 182 CCGCTCTCTTGATTCACCCCGGAGCTGCGCCACCATATCGCCCTGAGGCTGAGCT 241
QY 760 tctggg 765
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 242 TTTGGG 247

```

RESULT 13

```

AM456831/C      431 bp      mRNA      EST      24-FEB-2000
LOCUS      UI-M-BH3-agw-h-12-0-UI.s1 NIH.BMAP.M.54 Mus musculus cDNA clone
DEFINITION      UI-M-BH3-agw-h-12-0-UI 3', mRNA sequence.
ACCESSION      AM456831
VERSION      AM456831.1 GI:7027048
KEYWORDS
SOURCE
ORGANISM      EST.
Mus musculus
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 431)

```

```

REFERENCE
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
COMMENT
CONTACT: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mstemail.nih.gov

```

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized brain stems library cDNA library preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements. The following repetitive elements were found in this cDNA sequence: 18-106, /ORF1B/LTR/MaLR Seq primer: M13 Forward POLYA-yes.

FEATURES

SOURCE

```

Location/Qualifiers
1..431
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-agw-h-12-0-01"
/clone_1lb="NIH_BMAP_M.S4"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; site 1: Not I; site 2: Eco RI; The
NIH_BMAP_M.S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M.S4,
NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,
NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library
(NIH_BMAP_M.S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and
NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M.S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_1lb="NIH_BMAP_M.S4"
TAG_TISSUE="brain-stems"
TAG_SEQ="TCATC"
BASE COUNT      121 a      98 g      134 t
ORIGIN

```

Query Match 19.0%; Score 190; DB 10; Length 431;

Best Local Similarity 82.0%; Pred. No. 4.5e-45;

Matches 233; Conservative 0; Mismatches 45; Indels 6; Gaps 1;

```

QY 507 aagcacaaggaagtaactcagcaggaagtaaacctcgtctaccagagatgatggaag 566
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 431 AAGTCACGTGGAATATATACCCAGGAGTGAATCTCTTTACCGAGATGTTGTGAAG 372
QY 567 agttcctcattccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 626
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 371 AGTTCCTATTTCACAGTCACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 311
QY 627 aagtaagaagttttcaggaagaatcaagcctcctacacgtggtgaactctctgagc 666
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 312 -----AAGGTTTTCAGGAGATCAGAGTCATCCACCACTATGATGAATCTCTGAGCC 258

```

```

QY 687 caccgtactcgtcgtctacccctcgaatccaccocccaggttggtgacacagatgcc 746
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 257 CACATATACACACGCGCTATCTTGTGAAATTCACCCCGAGATCTGGAGCACCAATTCCTC 198
QY 747 tgaagatgaggttctggtctgcagcagcagcagcagcagcagcagcagcagcagc 790
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 197 TGAGCGTTGAGATTCGTAGATGTGAGCCGACGACGACGATCTGA 154

```

RESULT 14
AL552613 927 bp mRNA EST 16-FEB-2001
LOCUS AL552613 LIT.NR1006.PL2 Homo sapiens cDNA clone CSD01067G24 5
DEFINITION prime, mRNA sequence.
ACCESSION AL552613
VERSION AL552613.1 GI:12891683
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 927)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

```

Location/Qualifiers
1..927
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSD01067G24"
/clone_1lb="LIT.NR1006.PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@litech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      313 a      192 c      182 g      239 t      1 others
ORIGIN

```

Query Match 15.9%; Score 158.8; DB 10; Length 927;

Best Local Similarity 52.7%; Pred. No. 9.4e-36;

Matches 396; Conservative 0; Mismatches 347; Indels 9; Gaps 2;

```

QY 29 attatccagcgtcaagccagcagcagcagcagcagcagcagcagcagcagcagc 88
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 ATATATACAGGAGATCCAGCCAGGAGTCCAAACACTACTGTAAGTCTGCTATACCA 61
QY 89 cagttatcatcatgatgatgtcttcttggaagaagtgagcagactatcaggaatcc 148
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 GAGTCTATGATGCTTACAGGTCACCAACGATCAGTCTGCAATCTTCAAGAGAAACGC 121
QY 149 actggaacctaatgctctcttggcaatggtgatcctcgtggataaacaacatatt 208
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 122 ACAAGGAATGATGATGATTTTAAATGCAATTCAGATGCTCTACATATAAAGATCAG 181
QY 209 tttaaccctcaattctgctgataacatcgcttgacacacacattatagcttcgc 268
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 182 TTTCACCACTATTTGCTGATATATATATAGATCTCTCCAACTCGAGCCATATACGA 241
QY 269 agcactcttcgatgagttgagtggtgctgattaaatagttgagatgcatggga 328
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 242 CCAACCTTCGATTTGGAAGTCAAGGTTGTAGGTAATGATGATTTCCACACCCCTGGGT 301

```

OY	329	atgagagatgaagcaatcatcagatgcacagattactgctcattccattccattcaaatatg	388
Db	302	ATGGAATAATGGAAAGATATGAAACAAACAGCAAAATCACAGCTTCTTCGTTTAAGAATCTTGG	361
OY	389	t---ttgccacttgctctccttccaaaagctcgacttcaactcccaaggaggaaatgcc	445
Db	362	TGGGAGGATTTACTGGGAACCTTCCGCGCTGAAATCCCAAGGAGACTGTGAATGCC	421
OY	446	tggagacctcagtggaataatccaaaagatgctgtaagtggaacttcgaagaacaatg	505
Db	422	TGGCAAGCCAAAGGCAAAACACAAATATACAGAGGGCTAGAAATTGATCTACTCAAGATCAAG	481
OY	506	aaagtcacaagagtaacttactcagggaglaaaatctctgtcttcaaccgacatgataag	565
Db	482	AAGATTAACGGCAATTATATACACAGGGCTGCAAGCTCTGTGCTGTGAATATGTATGAAG	541
OY	566	gaagtctcaatctccagcagatcaagaatgagccataagtgagactctcttttccgaatgac	625
Db	542	AGCATATACCACTCCACTACATGACAGGAGGGAGTGGAAATGGAACCATACAGCTGAAATCC	601
OY	626	-----aaagtaaaagttttccagggaatcaagaacctcttcacacctgfygtaactct	679
Db	602	TCCATGTGTGACACAGATTTTGAAGGAAATACTATATCCAAAGACATGTGAAGAACTTT	661
OY	680	ctagaaccaccgttactgactcgcttaccttcgaattcaaccceccagatgfygfygacacag	739
Db	662	TTCAACCCCAATCATTTCCAGGTTTATCCGTCTCATTTCTTAACCATGTGAATCAAAAGT	721
OY	740	attgccctgagatgagagtgctcgtcgtcga	771
Db	722	ATTGACATTCCGCTGGAAGCTCTTTGGCTGTGA	753

RESULT	15
AL577036/c	
LOCUS	
DEFINITION	AL577036 745 bp mRNA EST 16-FEB-2001
ACCESSION	AL577036 LTL_NF1006.PL2 Homo sapiens cDNA clone CSDD1082Y022 3
VERSION	prime, mRNA sequence.
KEYWORDS	AL577036 AL577036.1 GI:12939773
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	I.(bases 1 to 745)
COMMENT	L.I.W.B., Gruber,C., Jessee,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr. Location/Qualifiers
FEATURES	

FEATURES	source	Location/Qualifiers
	1..745	
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="CSOD1082Y022"
		/clone_11b="L11_NFL006_P12"
		/tissue_type="Placenta"
		/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifestech.com URL: http://Fulllength.invitrogen.com "
BASE COUNT	180 a 160 c 158 g 238 t 9 others	
ORIGIN		

Query Match	13.78;	Score 136.6;	DB 10;	Length 745;
Best Local Similarity	52.18;	Pred. No. 3.2e-29;		
Matches 366;	Conservative 6;	Mismatches 321;	Indels 10;	Gaps 3;

Oy	79	ctcacctcttcagttatcatcactgatctagcttcgttcggagaagctgcacactatcgc	138
Db	743	CTTACACACAGAGTCTTATGACTTACAGTCCCAACACAGATCACTGGCAGATCTTCAA	688
Oy	139	aggaatctccactggaacacttaagtgtctctcttgagcaatgtagtcaatcctggataaa	196
Db	683	AGGGAACACACAGAGATGATGATATTTAATGGCAATTCAGATGCGCTCTACATTTAAA	624
Oy	199	acccaatatttttaacctcaattatgtctgcatacaccctgttcaccacaactcata	258
Db	623	AGGAATTCAGTTTGACCCACACCTATTGTGGCTAGATATATTAGATCTCTCCAACTCGAGC	566
Oy	259	taagacttcgaacactcttcgc-cataagatgtabgycgtgtgaatttaaatagtgcaga	317
Db	563	CTATTACAGACCTACCCCTTCGATGGGAATCGCAAGGTTGTGAGTAAATGATGTTCGA	504
Oy	318	tgcacattgggaatgtagagtaagaacaaatcagaatgcacagatctactgttctact	377
Db	503	CACCCCTGGGTATGAAAAATGAAAAAGTAAAAACAAGCAATACACAGCTTCTTGTTTA	444
Oy	378	ttaccaatatgt- - -ttgcacactgtctctctcctcaaaagctgcactccctcaaggga	433
Db	443	RGAAATCTTGGKGGGAGATTACTGGGAACCTTCGTCGCCGTGATGCCACGGAGAC	384
Oy	435	ggaagtaatgctctggaagactcaggttgataataccaagaagtgctgcagaaggtcttc	494
Db	383	GTTGTGATGCTGTGGCAAGCAAGCAAAACAACATAAGCATGCTGGCTRGAAATTTGATCAC	322
Oy	495	agaagaactgaaagtcacagagtaactcctcagagagtaaatcctcgtctacaaaga	555
Db	323	TCAAGATCARGAAGATAAACGGCAATTATAACACGGCGCTCAAGTCTGTCTCTGAAA	264
Oy	555	tgatgtgaagaagttcctcaatcctcagcagtcacaagtgcacatgagtgactctcttt	611
Db	263	TGTATGTAAAGACCTATACCATTCCTACTCACTGAGCGAGGAGTGGATGGAAACATATCA	201
Oy	615	ttcagaatggc- - - -aaagtaagaagttcttcagggaaatcaagaactcctccacactg	666
Db	203	GCGCTGAATCTCCATGCTGGGACATAGATTTTGAAGGAATTAATAATACCAAGGACATG	144
Oy	669	tgtatgaactctcagacccaccgtaactgactgcataccttgaattcaacccccagagt	722
Db	143	TGAAGACCTTTTCAACCCCCCAATCATTTTCCAGAGTTTATCCCTGTCATCTCTAAACAT	84
Oy	729	ggatgtcaccaagatgtccctgagatggaggtcttgggcttcga	771
Db	83	GGATTCAAAGTATGGACATGGCGCTGGAACTCHTTGGCTGTGA	41

Search completed: January 17, 2002, 14:59:01
Job time: 8662 sec

Tue Jan 22 15:12:49 2002

us-09-740-211-14_copy_4000_4999.rst

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:35:55 ; Search time 12.8 Seconds
(without alignments)
83.316 Million cell updates/sec

Title: US-09-740-211-15

Perfect score: 75

Sequence: 1 SFSQNPVYKRRHQR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_68:*

1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	86.7	2351	1 EZHU	coagulation factor
2	56	74.7	2319	2 A47004	coagulation factor
3	47	62.7	869	2 A25945	coagulation factor
4	47	62.7	2133	2 T27633	coagulation factor
5	42	56.0	1077	2 S70120	ATP-dependent RNA
6	41	54.7	2415	1 A33733	spectrin alpha cha
7	40	53.3	423	2 T40224	protein kinase - f
8	39	52.0	697	2 JN0602	rib protein kinase
9	39	52.0	3085	2 T00327	polyprotein - infe
10	39	52.0	3085	2 H81702	adherence factor T
11	38	50.7	199	1 S50398	ribosomal protein
12	38	50.7	199	2 S67618	ribosomal protein
13	38	50.7	269	2 T31318	hypothetical prote
14	38	50.7	303	2 T06981	low-molecularweig
15	38	50.7	304	2 T06505	glutinin low molec
16	38	50.7	332	2 G83066	conserved hypothet
17	38	50.7	347	2 S30921	ferredoxin-nitrit
18	38	50.7	436	2 F71358	hypothetical prote
19	38	50.7	604	2 T37994	probable splicing
20	38	50.7	623	2 T15510	hypothetical prote
21	38	50.7	1144	2 T27408	hypothetical prote
22	38	50.7	1307	2 T25563	hypothetical prote
23	38	50.7	1333	2 A57488	Kas guanine nucleo
24	37	49.3	356	2 S01992	glutinin low molec
25	37	49.3	359	2 T06982	glutinin low molec
26	37	49.3	374	2 T05923	glutinin low molec
27	37	49.3	403	2 B13738	probable recf prot
28	37	49.3	481	2 T84404	3-deoxy-D-manno-co
29	37	49.3	636	2 T38010	hypothetical prote

30	37	49.3	891	2 G84693	probable proline-r
31	37	49.3	1038	2 T15098	hypothetical prote
32	37	49.3	3255	2 G81702	adherence factor T
33	36	48.0	186	2 E72660	hypothetical prote
34	36	48.0	244	2 A72551	hypothetical prote
35	36	48.0	311	2 S57863	pyrroline-5-carbox
36	36	48.0	351	2 A34201	bone morphogenetic
37	36	48.0	366	2 F72703	probable heat choc
38	36	48.0	394	2 S45355	bone morphogenetic
39	36	48.0	433	2 T01574	indeterminate spik
40	36	48.0	561	2 S64068	hypothetical prote
41	36	48.0	607	2 F82502	GDEF family prote
42	36	48.0	633	2 B83478	probable two-compo
43	36	48.0	707	2 T26218	hypothetical prote
44	36	48.0	734	2 T13674	hypothetical prote
45	36	48.0	743	2 T13673	hypothetical prote

ALIGNMENTS

RESULT 1

EZHU

Coagulation factor VIII precursor [validated] - human

N:Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant co

C:Species: Homo sapiens (man)

C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Dec-2000

C/Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445;

R:Gitschler, J.; Wood, W.I.

Hum. Mol. Genet. 1, 199-200, 1992

A>Title: Sequence of the exon-containing regions of the human factor VIII gene.

A/Reference number: I54318; M01D:93265012

A/Accession: I54318

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1921; S', 1923-2351 <RES>

A/Cross-references: GB:M88648; NID:9182381; PIDN:AA52420.1; PID:9182383

R:Wood, M.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschler, J.; Keyt, B.; Seeb

Nature 312, 330-337, 1984

A>Title: Expression of active human factor VIII from recombinant DNA clones.

A/Reference number: A00525; M01D:85061548

A/Accession: A00525

A/Molecule type: mRNA

A/Residues: 1-2351 <MOO>

A/Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179

R:Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D

Nature 312, 342-347, 1984

A>Title: Molecular cloning of a cDNA encoding human antihemophilic factor.

A/Reference number: I58059; M01D:85061550

A/Accession: I58059

A/Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-74; V', 76-1259; E', 1261-2351 <RE2>

A/Cross-references: GB:K01740; NID:9182802; PIDN:AA52484.1; PID:9182803

R:Ruetz, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; K

B.; Randolph, A.; Urdes, M.S.; Valenzuela, P.; Dahl, H.H.; Pavalato, J.; Hansen, J.

DNA 4, 333-349, 1985

A>Title: Characterization of the polypeptide composition of human factor VIII:C and t

A/Reference number: A23584; M01D:86081164

A/Accession: A23584

A/Molecule type: mRNA

A/Residues: 1-2351 <TRD>

A/Cross-references: GB:M4113; NID:9182817; PIDN:AA52485.1; PID:9182818

R:Eaton, D.; Rodriguez, H.; Vohar, G.A.

Biochemistry 25, 505-512, 1986

A>Title: Proteolytic processing of human factor VIII. Correlation of specific cleavag

ity.

A/Reference number: A26174; M01D:86159740

A/Accession: A26174

A/Molecule type: Protein

A/Residues: 20-36; 392-399, 'X', 401-402; 1668-1678; 1709-1722, 'D', 1723-1725; 1741-1755 <EA

R:Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

Biochemistry 31, 3315-3325, 1992
 A:Title: Identification and functional importance of tyrosine sulfate residues within re
 A:Reference number: A42348; PMID:92207952
 A:Accession: A42348
 A:Molecule type: protein
 A:Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17
 A:Experimental source: recombinant material from Chinese hamster ovary cells
 A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
 R:Ray, P.J.; Smudzin, T.M.
 J. Biol. Chem. 264, 14005-14010, 1989
 A:Title: Intersubunit fluorescence energy transfer in human factor VIII.
 A:Reference number: A43986; PMID:89340500
 A:Accession: A43986
 A:Molecule type: protein
 A:Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <RAY>
 R:Heyte, A.; van Schijndel, H.B.; Niehs, C.; Huttnet, W.B.; Verbeet, M.P.; Mertens, K.;
 J. Biol. Chem. 266, 740-746, 1991
 A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
 A:Reference number: A56109; PMID:91093266
 A:Contents: annotation; sulfation
 R:Gitschier, J.; Wood, W.I.; Goralke, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vohar,
 Nature 312, 326-330, 1984
 A:Title: Characterization of the human factor VIII gene.
 A:Reference number: A56196; PMID:85061547
 A:Contents: annotation; introns
 R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
 Protein Sci. 4, 740-746, 1995
 A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains c
 A:Reference number: A56216; PMID:95338127
 A:Contents: annotation; disulfide bonds
 A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
 R:Fjalke, M.; Hedling, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
 Eur. J. Biochem. 234, 773-779, 1995
 A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
 A:Reference number: S63527; PMID:96163459
 A:Accession: S63527
 A:Molecule type: protein
 A:Residues: 733-752;753-759 <KJA>
 R:Lind, P.; Larsson, K.; Spita, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg
 Eur. J. Biochem. 232, 19-27, 1995
 A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
 A:Reference number: S66445; PMID:96048024
 A:Accession: S66445
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1668-1685 <LIN>
 C:Comment: Factor VIII is activated by factor xa and thrombin, but prolonged exposure pr
 C:Genetics:
 A:Gene: GDB:F8C
 A:Cross-References: GDB:119124; OMIM:306700
 A:Map position: Xq28-Xq28
 A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
 C:Function:
 A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa prc
 A:Pathway: blood coagulation
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:20-235/Product: coagulation factor VIII #status experimental <Mat>
 F:20-740/Product: coagulation factor VIIIA heavy chain #status experimental <ACH>
 F:20-356/Domains: A1 <DNI>
 F:23-348/Domains: ferroxidase repeat homology <FO1>
 F:392-759/Domains: A2 <DA2>
 F:402-730/Domains: ferroxidase repeat homology <FO2>
 F:760-1667/Domains: B <DBO>
 F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <ACL>
 F:1709-2038/Domains: A3 <DA3>
 F:1716-2038/Domains: ferroxidase repeat homology <FO3>
 F:2039-2191/Domains: C1 <DC1>
 F:2039-2188/Domains: discoidin I amino-terminal homology <DN1>
 F:2192-2351/Domains: C2 <DC2>
 F:2192-2345/Domains: discoidin I amino-terminal homology <DN2>
 F:60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,11

F:172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #sta
 F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted
 F:365-737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experiment
 F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
 F:414,426/Binding site: sulfate (Tyr) (covalent) #status predicted
 F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
 F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
 F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
 F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa, thrombin) #status experime
 F:2193-2345/Disulfide bonds: #status predicted

Query Match 86.7%; Score 65; DB 1; Length 2351;
 Best Local Similarity 100.0%; Pred. No. 0.0076;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 3 SONPPVLRKHQR 14
 DB 1656 SONPPVLRKHQR 1667

RESULT 2
 A47004
 C:Accession: A47004
 C:Species: Mus musculus (house mouse)
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999
 R:Elider, B.; Jakich, D.; Gitschier, J.
 Genomics 16, 374-379, 1993
 A:Title: Sequence of the murine factor VIII cDNA.
 A:Reference number: A47004; PMID:93300511
 A:Accession: A47004
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2319 <ELD>

A:Cross-References: GB:I05573; NID:q192456; PIDN:AAA37385.1; PID:q192457
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:23-349/Domains: ferroxidase repeat homology <FO1>
 F:402-730/Domains: ferroxidase repeat homology <FO2>
 F:1666-2006/Domains: ferroxidase repeat homology <FO3>
 F:2007-2156/Domains: discoidin I amino-terminal homology <DN1>
 F:2160-2313/Domains: discoidin I amino-terminal homology <DN2>

Query Match 74.7%; Score 56; DB 2; Length 2319;
 Best Local Similarity 91.7%; Pred. No. 0.28;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Query 3 SONPPVLRKHQR 14
 DB 1629 SQRPVLRKHQR 1640

RESULT 3
 A25945
 C:Accession: A25945
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 21-Jul-2000
 R:Toole, J.J.; Pittman, D.D.; Orr, E.C.; Murtha, P.; Wasley, L.C.; Kaufman, R.J.
 Proc. Natl. Acad. Sci. U.S.A. 83, 5939-5942, 1986
 A:Title: A large region (approx 195 kDa) of human factor VIII is dispensable for in v
 A:Reference number: A25945; PMID:86287369
 A:Accession: A25945
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-869 <TOO>
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid

Query Match 62.7%; Score 47; DB 2; Length 869;
 Best Local Similarity 88.9%; Pred. No. 3.6;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PPVLRKHQR 14
|||||

Db 737 PPVLRKHQR 745

RESULT 4

T42763
coagulation factor VIII precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
C:Accession: T42763
R:Lollar, P.
submitted to the EMBL Data Library, August 1996
A:Reference number: 222269
A:Accession: T42763
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2133 <LOL>
A:Cross-references: EMBL:U49517; NID:g1511633; PID:AAB06705.1
C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
F:23-349/Domain: ferroxidase repeat homology <FOX1>
F:402-730/Domain: ferroxidase repeat homology <FOX2>
F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match Best Local Similarity 62.7%; Score 47; DB 2; Length 2133;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PPVLRKHQR 14
|||||

Db 1441 PPVLRKHQR 1449

RESULT 5

S70120
ATP-dependent RNA helicase homolog YDR291w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D9819.1
C:Species: Saccharomyces cerevisiae
C>Date: 24-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 02-Feb-2001
C:Accession: S70120
R:Fulton, L.
submitted to the EMBL Data Library, May 1996
A:Description: The sequence of S. cerevisiae cosmid 9819.
A:Reference number: S70114
A:Accession: S70120
A:Molecule type: DNA
A:Residues: 1-1077 <FUL>
A:Cross-references: EMBL:U51031; NID:g1332635; PID:AAB64466.1; PID:g1230655; MIPS:YDR29
C:Genetics:
A:Map position: 4R
A>Note: YDR291w
C:Keywords: ATP; nucleotide binding; P-loop
F:312-319/Region: nucleotide-binding motif A (P-loop)
F:419-424/Region: nucleotide-binding motif B
F:423-426/Region: DEXH motif

Query Match Best Local Similarity 56.0%; Score 42; DB 2; Length 1077;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPPVLRKHQR 14
|||||

Db 500 NPPVLRKHQR 509

RESULT 6

A33733
spectrin alpha chain - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: A33733; A49468; B49468
R:Dubreuil, R.R.; Byers, T.J.; Sillman, A.L.; Bar-Zvi, D.; Goldstein, L.S.B.; Branton
J. Cell Biol. 109, 2197-2205, 1989

A:Title: The complete sequence of Drosophila alpha-spectrin: conservation of structure
A:Reference number: A33733; MUID:90037215

A:Accession: A33733
A:Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-2415 <DUB>

A:Cross-references: GB:M26400; NID:g158488; PID:AAA28907.1; PID:g158489
R:Lee, J.K.; Coyne, R.S.; Dubreuil, R.R.; Goldstein, L.S.; Branton, D.

A:Title: Cell shape and interaction defects in alpha-spectrin mutants of Drosophila m
A:Reference number: A49468; MUID:94103334

A:Accession: A49468
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-150 <LEE>

A:Cross-references: GB:S67765; NID:g544666; PID:AAB29441.1; PID:g544667
A:Note: sequence extracted from NCBI backbone (NCBIN:141786, NCBIN:141790, NCBI:1417

A:Accession: B49468
A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA
A:Residues: 2192-2415 <LEE>

A:Cross-references: GB:S67765; NID:g544666; PID:AAB29442.1; PID:g544668
A:Note: sequence extracted from NCBI backbone (NCBI:141794)

C:Genetics:
A:Gene: FlyBase:fbagr

A:Cross-references: FlyBase:FBgn0003470
C:Superfamily: spectrin alpha chain; calmodulin repeat homology; SH3 homology; spectr
C:Keywords: actin binding; cytoskeleton; EF hand

F:46-150/Domain: spectrin/dystrophin repeat homology <SP1>
F:151-256/Domain: spectrin/dystrophin repeat homology <SP2>
F:257-362/Domain: spectrin/dystrophin repeat homology <SP3>
F:363-468/Domain: spectrin/dystrophin repeat homology <SP4>
F:469-574/Domain: spectrin/dystrophin repeat homology <SP5>
F:575-679/Domain: spectrin/dystrophin repeat homology <SP6>
F:680-785/Domain: spectrin/dystrophin repeat homology <SP7>
F:786-891/Domain: spectrin/dystrophin repeat homology <SP8>
F:892-976/Domain: spectrin/dystrophin repeat homology #status atypical <SP9>
F:977-1024/Domain: SH3 homology <SH3>
F:1072-1179/Domain: spectrin/dystrophin repeat homology <SP10>
F:1180-1285/Domain: spectrin/dystrophin repeat homology <SP11>
F:1286-1391/Domain: spectrin/dystrophin repeat homology <SP12>
F:1392-1497/Domain: spectrin/dystrophin repeat homology <SP13>
F:1498-1604/Domain: spectrin/dystrophin repeat homology <SP14>
F:1605-1710/Domain: spectrin/dystrophin repeat homology <SP15>
F:1711-1816/Domain: spectrin/dystrophin repeat homology <SP16>
F:1817-1922/Domain: spectrin/dystrophin repeat homology <SP17>
F:1923-2029/Domain: spectrin/dystrophin repeat homology <SP18>
F:2038-2143/Domain: spectrin/dystrophin repeat homology <SP19>
F:2152-2252/Domain: spectrin/dystrophin repeat homology <SP20>
F:2265-2297/Domain: calmodulin repeat homology <EF1>
F:2308-2340/Domain: calmodulin repeat homology <EF2>

Query Match Best Local Similarity 54.7%; Score 41; DB 1; Length 2415;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SFSQNPVLRKH 12
|:|:|:|

Db 1995 SHAQSPVLRKH 2006

RESULT 7

T40224
protein kinase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: T40224
 R:Moore, V.; Rajandream, M. A.; Barrell, B. G.; Badcock, K.; Churcher, C. M.
 Submitted to the EMBL Data Library, March 1998
 A:Reference number: Z21914
 A:Accession: T40224
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-423 <MO>
 A:Cross-references: EMBL:AL022172; PIDN:CAA18163.1; GSPDB:GN00067; SPDB:SPBC32C12.03C
 A:Experimental source: strain 97zh-; cosmid c32C12
 C:Genetics:
 A:Gene: SPDB:SPBC32C12.03C
 A:Map position: 2
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 53.3%; Score 40; DB 2; Length 423;
 Best Local Similarity 77.8%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 PVLKRHR 14
 ||| |||
 Db 393 PPVSRHR 401

RESULT 8
 rfb protein - Escherichia coli (strain B41)
 N:Alternate names: protein B
 C:Species: Escherichia coli
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 08-Oct-1999
 C:Accession: JN0602; S33065
 R:Cheah, K. C.; Manning, P. A.
 Gene 123, 9-15, 1993
 A>Title: Inactivation of the Escherichia coli B41 (O101:K99/F41) rfb gene encoding an 80
 A:Reference number: JN0602; MUID:93138438
 A:Accession: JN0602
 A:Molecule type: DNA
 A:Residues: 1-697 <CHE>
 A:Cross-references: EMBL:X59852; NID:942718; PIDN:CAA42515.1; PID:942719
 C:Comment: This protein is located in the cytoplasmic membrane and has a role in O-anti-
 C:Genetics:
 A:Gene: rfb
 C:Keywords: membrane protein

Query Match 52.0%; Score 39; DB 2; Length 697;
 Best Local Similarity 61.5%; Pred. No. 71;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 FSQNPVLKRHR 14
 ||| |||
 Db 684 FGQVPRNLKRRHR 696

RESULT 9
 T00327
 polyprotein - infectious flacherie virus
 C:Species: infectious flacherie virus
 C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
 C:Accession: T00327
 R:Issawa, H.; Asano, S.; Sahara, K.; Iizuka, T.; Bando, H.
 Arch. Virol. 143, 127-143, 1998
 A>Title: Analysis of genetic information of an insect picorna-like virus, infectious fla-
 ke) viruses.
 A:Reference number: Z14139; MUID:98168871
 A:Accession: T00327
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: genomic RNA
 A:Residues: 1-3085 <ISA>
 A:Cross-references: EMBL:AB000906; NID:93025414; PIDN:BA25371.1; PID:93025415
 C:Keywords: polyprotein

Query Match 52.0%; Score 39; DB 2; Length 3085;
 Best Local Similarity 54.5%; Pred. No. 3,4e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 FSQNPVLKRHR 12
 ||| |||
 Db 2070 FNQNSPIKRRH 2080

RESULT 10
 H81702
 adherence factor TC0438 [imported] - Chlamydia muridarum (strain Nig9)
 C:Species: Chlamydia muridarum; Chlamydia trachomatis Mofn
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
 C:Accession: H81702
 R:Read, T. D.; Brunham, R. C.; Shen, C.; Gill, S. R.; Heidelberg, J. F.; White, O.; Hicke
 C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salbe
 Nucleic Acids Res. 28, 1397-1406, 2000
 A>Title: Genome sequences of Chlamydia trachomatis Mofn and Chlamydia pneumoniae AR39
 A:Reference number: A81500; MUID:20150255
 A:Accession: H81702
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-3335 <TET>
 A:Cross-references: GB:AE002311; GB:AE002160; NID:97190477; PIDN:AAF39292.1; PID:9719
 A:Experimental source: strain Nig9 (Mofn)
 C:Genetics:
 A:Gene: TC0438

Query Match 52.0%; Score 39; DB 2; Length 3335;
 Best Local Similarity 87.5%; Pred. No. 3,7e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVLK 10
 ||| |||
 Db 2174 SQNPVLK 2181

RESULT 11
 S50398
 ribosomal protein L13.e-B, cytosolic - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YMO375.1lc; protein YML142C
 C:Species: Saccharomyces cerevisiae
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 22-Oct-1999
 C:Accession: S50398
 R:Badcock, K.; Churcher, C.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: S50388
 A:Accession: S50398
 A:Molecule type: DNA
 A:Residues: 1-199 <BAD>
 A:Cross-references: EMBL:Z47071; NID:9606429; PIDN:CAA87356.1; PID:9606440; GSPDB:GNO
 C:Genetics:
 A:Gene: MTPS:YMR142C
 A:Map position: 13R
 A:Insertions: 2/1
 C:Superfamily: rat ribosomal protein L13
 C:Keywords: cytosol; protein biosynthesis; ribosome

Query Match 50.7%; Score 38; DB 1; Length 199;
 Best Local Similarity 50.0%; Pred. No. 29;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 SFSQNPVLKRHR 14
 :||| |||
 Db 2 AISKNPILKRRHR 15

RESULT 12

Query Match	50.7%;	Score 38;	DB 2;	Length 199;
Best Local Similarity	50.0%;	Pred. No. 29;		
Matches	7;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0;
QY	1	SFSQSNPVYLKRROR	14	
	:	: : : : :		
Db	2	AIRKNPILIKNHR	15	

Query Match	50.7%	Score 38;	DB 2;	Length 269;
Best local Similarity	63.6%	Pred. No. 39;		
Matches	7;	Conservative	2;	Mismatches
			2;	Indels
			0;	Gaps
QY	4 QNPVPLKRRHR	14		
	:::			
Db	50 QNPVILRRGR	60		

Query Match	50.7%;	Score 38;	DB 2;	Length 303;
Best Local Similarity	50.0%;	Pred. No. 45;		
Matches	7;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0;
QY	1 SFSQNPVYLKRHR	14		
Db	43 SFSQNPVYLSQDQDQ	56		

Query Match	50.7%	Score 38;	DB 2;	length 304;
Best Local Similarity	50.0%;	Pred. No. 45;		
Matches	7;	Conservative	3;	Mismatches 4;
				Indels 0;
				Gaps 0;
QY	1	SFSQNPVLRHQR	14	
			:::	:
Db	43	SFSQNPFSQSQQQQ	56	

Search completed: January 17, 2002, 12:37:06
Job time: 71 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:35:55 ; Search time 23.64 Seconds

(without alignments)
43,867 Million cell updates/sec

Title: US-09-740-211-15

Perfect score: 75
Sequence: 1 SFSQNPVLRHQR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	14	AAW04651	Peptide linker for
2	75	100.0	14	AAE04332	Human factor VIII
3	75	100.0	1438	AAE01252	B-domain deleted f
4	75	100.0	1440	AAE01252	Factor VIII:SQ
5	75	100.0	1457	AAW46246	Human factor VIII
6	75	100.0	1457	AAW44372	Human factor VIII
7	75	100.0	1457	AAW21675	Beta-domain delete
8	75	100.0	36	AAE56776	Human anti-haemoph
9	65	86.7	86	AAE56775	Human anti-haemoph
10	65	86.7	211	AAE50091	Truncated derivati
11	65	86.7	211	AAE07203	Human factor VIII

12	65	86.7	1283	21	AAE07205	Human factor VIII
13	65	86.7	1284	6	AAE50108	Factor-VIII deriv
14	65	86.7	1516	9	AAE08265	Modified factor VI
15	65	86.7	1661	18	AAE18670	Factor VIII-dB695-
16	65	86.7	2098	17	AAE86863	Factor-VIII. Homo
17	65	86.7	2332	8	AAE71727	Factor VIII:c varia
18	65	86.7	2332	8	AAE71728	Factor VIII:c varia
19	65	86.7	2332	8	AAE71729	Factor VIII:c varia
20	65	86.7	2332	14	AAE43257	Human factor VIII
21	65	86.7	2332	18	AAE33222	Procoagulant-activ
22	65	86.7	2332	18	AAE33223	Procoagulant-activ
23	65	86.7	2332	18	AAE33224	Procoagulant-activ
24	65	86.7	2332	18	AAE33225	Procoagulant-activ
25	65	86.7	2332	18	AAE33226	Procoagulant-activ
26	65	86.7	2332	19	AAE53483	Human factor VIII
27	65	86.7	2332	19	AAE44132	Homo sapiens modif
28	65	86.7	2332	20	AAE31594	Human factor VIII
29	65	86.7	2332	21	AAE57847	Human factor VIII
30	65	86.7	2332	22	AAE71902	N-terminal truncat
31	65	86.7	2332	22	AAE50465	Human factor VIII
32	65	86.7	2332	18	AAE14422	Active factor VIII
33	65	86.7	2342	18	AAE11349	Active factor VIII
34	65	86.7	2343	18	AAE11393	Active factor VIII
35	65	86.7	2343	18	AAE11335	Active factor VIII
36	65	86.7	2344	18	AAE11432	Active factor VIII
37	65	86.7	2344	18	AAE11410	Active factor VIII
38	65	86.7	2344	18	AAE11384	Active factor VIII
39	65	86.7	2344	18	AAE11361	Active factor VIII
40	65	86.7	2344	18	AAE11367	Active factor VIII
41	65	86.7	2344	18	AAE11370	Active factor VIII
42	65	86.7	2344	18	AAE11331	Active factor VIII
43	65	86.7	2345	18	AAE11466	Active factor VIII
44	65	86.7	2345	18	AAE11415	Active factor VIII
45	65	86.7	2345	18	AAE11403	Active factor VIII

ALIGNMENTS

RESULT 1	AAW04651	standard; peptide; 14 AA.
ID	AAW04651	
AC	AAW04651	
XX		
DT	06-AUG-1997	(first entry)
XX		
DE		Peptide linker for truncated factor VIII 90 and 80 kd fragments.
XX		
KW		linker; cell culture medium; liposome; lipid; phosphatidylcholine;
KW		phosphatidylserine; increase; production; recombinant factor VIII;
KW		truncated; haemophilia; treatment; plasma protein.
XX		
OS	Synthetic.	
XX		
PN	EP745672-A2.	
XX		
PD	04-DEC-1996.	
XX		
PF	25-APR-1996;	96EP-0106482.
XX		
PR	17-APR-1996;	96US-0634001.
XX	04-MAY-1995;	95US-0434900.
PA	(FARB) BAYER CORP.	
PA	(CHAN/) CHAN S.	
PA	(MILE) MILES INC.	
XX		
PI	Chan S;	
XX		
DR	WPI; 1997-013695/02.	
XX		
PT	Culture medium contg. liposome like substance comprising at least 2	

PT different lipid(s) - useful for increasing prodn. of recombinant
 PT Factor VIII in mammalian cell cultures
 XX
 XX
 PS Claim 15; Page 7; 7pp; English.

CC A novel cell culture medium contg. a liposome like substance, comprises
 CC at least 2 different lipids in a molar ratio sufficient to assure a
 CC 4-fold increase in factor VIII expression in a mammalian cell culture
 CC system. The culture medium is useful to increase production of
 CC recombinant factor VIII 4-fold and truncated factor VIII 3-fold.
 CC Factor VIII is a plasma protein required for clotting of the blood,
 CC useful in treatment of haemophilia. Liposome like substances contg.
 CC lipids such as phosphatidylcholine (PC), phosphatidylethanolamine (PE)
 CC or phosphatidylserine (PS) alone have no effect on recombinant factor
 CC VIII expression in BHK-21 and 293S cells. The present sequence is a
 CC linker joining truncated recombinant factor VIII 90 and 80 kd
 CC fragments.

CC Sequence 14 AA;

Query Match 100.0%; Score 75; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.5e-07;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVLRKHQR 14
 ||||||||||||
 Db 1 sfsqnpvlkrhqr 14

RESULT 2

AAE04332
 ID AAE04332 standard; peptide; 14 AA.

AC AAE04332;

DT 04-SEP-2001 (first entry)

DE Human factor VIII B-domain partial sequence.

XX Recombinant adeno-associated viral vector; rAAV; factor VIII; FVIII;
 KW blood clotting disorder; gene therapy; haemophilia A; human; B-domain.
 XX

OS Homo sapiens.

XX WO200145510-A1.

XX 28-JUN-2001.

PF 21-DEC-2000; 2000WO-US34925.

PR 22-DEC-1999; 99US-0470618.

XX (AVIG-) AVIGEN INC.

PI Couto LB, Colosi PC, Qian X;

XX WPI; 2001-417955/44.

PT Treating blood clotting disorder, especially hemophilia in mammals, by
 PT administering recombinant adeno-associated vectors which express blood
 PT coagulation factor VIII -

PS Claim 6; Fig 2; 90pp; English.

XX The present invention relates to a method for treating a subject
 CC suffering from a blood clotting disorder. The method comprises
 CC administering a recombinant adeno-associated virion (rAAV) comprising
 CC a nucleotide sequence encoding the light chain of factor VIII and a
 CC second recombinant adeno-associated virion comprising a nucleotide
 CC sequence encoding the heavy chain of factor VIII. The rAAV vector is
 CC useful in gene therapy for treating haemophilia A in mammals, in
 CC particular humans. The rAAV vector provides high level and long term

CC expression of biologically active clotting factor VIII in vivo.
 CC The present sequence represents partial human factor VIII B-domain.
 CC This sequence is encoded by rAAV vectors of the invention.

XX Sequence 14 AA;

Query Match 100.0%; Score 75; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.5e-07;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVLRKHQR 14
 ||||||||||||
 Db 1 sfsqnpvlkrhqr 14

RESULT 3

AAE01262
 ID AAE01262 standard; protein; 1438 AA.

AC AAE01262;

DT 25-SEP-2000 (first entry)

DE B-domain deleted factor VIII sequence.

XX Factor VIII; procoagulant; adenovirus; adeno-associated strain;
 KW gene therapy; human Burkitt's lymphoma; HKB; therapy;
 KW therapeutic protein; vector; Epstein-Barr virus; human.
 XX

OS Homo sapiens.

XX WO200034505-A1.

XX 15-JUN-2000.

PF 08-DEC-1999; 99WO-US29169.

PR 10-DEC-1998; 98US-0209916.

XX (FARB) BAYER CORP.

PI Cho M, Chan SY, Kelsey W, Yee H;

XX WPI; 2000-431311/37.

PT Producing cells expressing a protein having factor VIII procoagulant
 PT activity especially, human factor VIII in an industrial scale. Involves
 PT expressing a vector comprising a sequence coding for factor VIII in
 PT human cells

PS Claim 7; Fig 1; 27pp; English.

XX Producing cells expressing a protein having factor VIII procoagulant
 CC activity, comprises contacting the cells with a vector comprising a
 CC selectable marker and a sequence coding for the protein having factor
 CC VIII procoagulant activity operably linked to a promoter. The cells
 CC are then selected and individual clones expressing high levels of the
 CC protein are isolated from the selected cells. The cells produced by
 CC the method are not only useful for producing protein having factor
 CC VIII procoagulant activity but also for producing adenovirus and
 CC adeno-associated virus strains for gene therapy. The advantage of
 CC having cells producing protein with factor VIII procoagulant activity
 CC is that factor VIII protein can be produced on an industrial scale
 CC in the range of 2-4 pg/cell/day. Human Burkitt's lymphoma (HKB)
 CC cells provide a protein-free production system to produce not only
 CC B-domain deleted factor VIII but also other therapeutic proteins. The
 CC vector used in the method preferably comprises B-domain deleted
 CC factor VIII (BDD-FVIII), a transcriptional unit for BDD-FVIII and a
 CC selectable marker, dihydrofolate reductase (dhfr). In addition, a
 CC terminal repeat sequence from Epstein-Barr virus is inserted into the
 CC vector to increase integration efficiency.

SQ Sequence 1438 AA;

Query Match 100.0%; Score 75; DB 21; Length 1438;

Best Local Similarity 100.0%; Pred. No. 6.8e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVPLKRHR 14

Db 741 sfsqnpvplkrhr 754

RESULT 4

ID AAR12971 standard; protein; 1440 AA.

AC AAR12971;

DE 02-OCF-1991 (first entry)

XX Factor VIII:SQ.

XX Factor VIII; B domain; haemophilia.

XX MO9109122-A.

XX 27-JUN-1991.

XX 06-DEC-1990; 90WO-SE00809.

XX 15-DEC-1989; 89SE-0004239.

XX (KABI) KABIYTRDM AB.

XX Almstedt AB, Hellstrom EM, Larsson K, Lind P, Sandberg HI, Spira J;

XX Sydow-Backman M;

XX WPI: 1991-208148/28.

XX Recombinant human factor VIII deriv. deoxyribonucleic acid -

XX encoding protein comprising two chains linked by segment of B domain.

XX Disclosure; Fig 1; 35pp; English.

XX The protein is a fusion between Phe 742 and Ser 1637 of the factor

XX VIII protein (factor VIII:SQ). In order to produce a

XX factor VIII deletion derivative that can be produced in vivo and/or

XX in vitro, to a two chain protein consisting of polypeptide chains of

XX 90 kD and 80 kD, the amino acid sequences surrounding Arg 740 and

XX Arg 1648 have to be conserved in order to preserve the structural

XX requirements for correct cleavage. In this example, amino acids 743

XX to 1636 of the full-length factor VIII polypeptide are deleted. A

XX new polypeptide chain is obtd. where there are 14 amino acids

XX linking Arg 740 and Arg 1648. Of these 14 amino acids, the sequence

XX of the five N-terminal ones directly corresponds to the five amino

XX acids following Arg 740 in full-length factor VIII. Also, the sequence

XX of the 12 C-terminal amino acids of the above 14 amino acids fragment

XX directly corresponds to the 12 amino acids preceding Glu 1689 in full-

XX length factor VIII, thus creating a 3 amino acid overlap between the

XX N- and C-terminal regions of the B-domain.

XX The factor VIII deriv. is useful for treating haemophilia or

XX haemophilia A. It has the biological characteristics of plasma derived

XX factor VIII.

XX In order to index this example, the factor VIII:QD amino acid

XX sequence was retrieved from WO8800831 (AAB80265).

XX The amino acid numbering in the above comments is reproduced from the

XX fig. description in the specification. Note that Arg 740 is Arg 742

XX in the indexed sequence, etc. Also, Asn 745 (N-terminal link overlap)

XX in Asp 747 in AAB80265, but indexed as Asn to reproduce the fusion

XX fragment as shown in fig.1, and Asn 745 (C-terminal link overlap).

XX Sequence 1440 AA;

Query Match 100.0%; Score 75; DB 12; Length 1440;

Best Local Similarity 100.0%; Pred. No. 6.8e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVPLKRHR 14

Db 743 sfsqnpvplkrhr 756

RESULT 5

ID AAM46246 standard; protein; 1457 AA.

AC AAM46246;

DE 06-AUG-1998 (first entry)

XX Human factor VIII beta-domain deleted SQN deletion protein sequence.

XX Replication defective; recombinant retrovirus; RRV; therapeutic protein;

XX haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;

XX hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;

XX cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;

XX hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;

XX Guacher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;

XX inflammatory disease; factor VIII.

XX Homo sapiens.

XX WO9800541-A2.

XX 08-JAN-1998.

XX 02-JUL-1997; 97WO-US11784.

XX 04-JUN-1997; 97US-0869309.

XX 03-JUL-1996; 96US-0645601.

XX 13-AUG-1996; 96US-0696381.

XX (CHIR) CHIRON CORP.

XX Allen JR, Barber JR, Boder M, Chang SMW, Chong K;

XX De LA VEGA D, Depoloni, Greengard J, Hsu DC, Ibanez CE;

XX Jolly DJ, Lee R, Mittelstaedt DM, Prusek CE, Respass JG;

XX WPI: 1998-086966/08.

XX N-PSDB; AAV19581.

XX New replication defective recombinant retro-virus - which can be

XX administered to provide long term systemic expression of therapeutic

XX protein in blood, useful in, e.g. treating hyper-coagulable

XX disorders

XX Example 28; Pages 213-217; 272pp; English.

XX This is the beta-domain deleted SQN deletion protein of human factor

XX VIII. The encoding DNA is used to construct recombinant retroviral

XX vectors expressing human factor VIII. The invention provides the

XX preparation of replication defective recombinant retrovirus (RRV)

XX expressing a therapeutic protein. The RRV preparation is resistant to

XX degradation by human complement and is capable of inducing long term

XX systemic expression of the therapeutic protein when administered

XX intravenously to a human. The long term systemic expression results in a

XX measurable level of the therapeutic protein being produced in the blood

XX of the human for a period of at least 30 days after the administration of

XX the RRV vector preparation. RRV's can be used for in vivo delivery of

XX therapeutic protein to treat, e.g. haemophilia A, haemophilia B,

XX thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,

XX disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,

XX severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's

XX muscular dystrophy, inherited emphysema, familial hypercholesterolemia,

XX diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-

OY 1 SFSQNPVLKRHR 14
 DB 760 SFSQNPVLKRHR 773

RESULT 8

AAR56776
 ID AAR56776 standard; peptide; 36 AA.

AC AAR56776;

DE 03-APR-1995 (first entry)

XX Human anti-haemophilic factor VIII B domain C-terminal fragment.

XX Antihaemophilic factor VIII; B domain; monoclonal antibody; Mab;

OS Homo sapiens.

PN JP06205696-A.

PD 26-JUL-1994.

PF 11-JAN-1993; 93JP-0002537.

PR 11-JAN-1993; 93JP-0002537.

PA (TEIJ) TEIJIN LTD.

PA (KAGA) ZH KAGAKU & KESSEI RYOHO KENKYUSHO.

DR WPI; 1994-275529/34.

XX Monoclonal antibody binding to unnatural human anti-haemophilic

XX factor VIII B domain - is used in immuno-adsorption process to

XX bind impurities

XX Claim 3; Page 10; 12pp; Japanese.

XX A monoclonal antibody (Mab) directed against this peptide may be

XX used in an immuno-adsorption process to remove unnatural type human

XX antihaemophilic factor VIII L chain having this peptide fused to its

XX C-terminus (an impurity on preparation of the human antihaemophilic

XX factor VIII protein complex). The Mab bonds to unnatural type human

XX antihaemophilic factor VIII L chain fused with the peptide but does

XX not bond substantially to natural type human antihaemophilic factor

XX VIII L chain. This peptide corresponds to amino acids 1614-1649 of

XX human antihaemophilic factor VIII-B domain.

XX Sequence 36 AA;

XX

XX

XX

XX

XX

XX

XX Homo sapiens.

XX JP06205696-A.

XX 26-JUL-1994.

XX 11-JAN-1993; 93JP-0002537.

XX 11-JAN-1993; 93JP-0002537.

XX (TEIJ) TEIJIN LTD.

XX (KAGA) ZH KAGAKU & KESSEI RYOHO KENKYUSHO.

XX WPI; 1994-275529/34.

XX Monoclonal antibody binding to unnatural human anti-haemophilic

XX factor VIII B domain - is used in immuno-adsorption process to

XX bind impurities

XX Claim 3; Page 10; 12pp; Japanese.

XX A monoclonal antibody (Mab) directed against this peptide may be

XX used in an immuno-adsorption process to remove unnatural type human

XX antihaemophilic factor VIII L chain having this peptide fused to its

XX C-terminus (an impurity on preparation of the human antihaemophilic

XX factor VIII protein complex). The Mab bonds to unnatural type human

XX antihaemophilic factor VIII L chain fused with the peptide but does

XX not bond substantially to natural type human antihaemophilic factor

XX VIII L chain. This peptide corresponds to the 86 C-terminal amino

XX acids of human antihaemophilic factor VIII-B domain (amino acids

XX 1563-1648 of the mature protein).

XX Sequence 86 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Homo sapiens.

XX JP06205696-A.

XX 26-JUL-1994.

XX 11-JAN-1993; 93JP-0002537.

XX 11-JAN-1993; 93JP-0002537.

XX (TEIJ) TEIJIN LTD.

XX (KAGA) ZH KAGAKU & KESSEI RYOHO KENKYUSHO.

XX WPI; 1994-275529/34.

XX Monoclonal antibody binding to unnatural human anti-haemophilic

XX factor VIII B domain - is used in immuno-adsorption process to

XX bind impurities

XX Claim 3; Page 10; 12pp; Japanese.

XX A monoclonal antibody (Mab) directed against this peptide may be

XX used in an immuno-adsorption process to remove unnatural type human

XX antihaemophilic factor VIII L chain having this peptide fused to its

XX C-terminus (an impurity on preparation of the human antihaemophilic

XX factor VIII protein complex). The Mab bonds to unnatural type human

XX antihaemophilic factor VIII L chain fused with the peptide but does

XX not bond substantially to natural type human antihaemophilic factor

XX VIII L chain. This peptide corresponds to the 86 C-terminal amino

XX acids of human antihaemophilic factor VIII-B domain (amino acids

XX 1563-1648 of the mature protein).

XX Sequence 86 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

XX DR WPI; 1985-191681/32.
XX DR N-PSDB; AAN50105.
XX PT Prodn. of human Factor VIII C or its precursors or sub-units - by
XX PT using recombinant DNA techniques with Factor VIII C gene
XX PT expressed in host.
XX PS Disclosure; Page 32; 58pp; English.
XX CC The sequence encodes a truncated derivative of human Factor-VIIIc,
XX CC which is useful as an anticoagulant for the treatment of
XX CC haemophilia.
XX SQ Sequence 211 AA:

Query Match      86.7%; Score 65; DB 6; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.00055;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVYLRKHQR 14
   |||||
DB 115 sqnpvylkrhqr 126

RESULT 11
AAB07203
XX ID AAB07203 standard; Protein; 211 AA.
XX AC AAB07203;
XX DT 11-OCT-2000 (first entry)
XX DE Human Factor VIIIc 77/80kd subunit.
XX KW Factor VIII C domain; human; blood clotting; haemophilia.
XX OS Homo sapiens.
XX PA (CHIR ) CHIRON CORP.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Kuo G, Rasmussen ME, Maslarsz FR, Valenzuela P, Truett M;
XX PI Favalloro J;
XX DR WPI; 2000-367967/32.
XX DR N-PSDB; AAA58440.
XX PT DNA sequence of portion of human Factor VIIIc for treating and
XX PT preventing the symptoms of haemophilia -
XX PS Disclosure; Page 16; 39pp; English.
XX CC The present sequence is the protein sequence for the 77/80kd subunit
XX CC of human Factor VIIIc. The sequence was identified by creating and
XX CC sequencing a genomic DNA library of the Factor VIIIc gene. Factor VIIIc
XX CC is a plasma protein involved in blood coagulation, and is absent or
XX CC defective in haemophilia A. The Factor VIIIc protein can, therefore, be
XX CC used to treat haemophilia, as well as in the production of monoclonal
XX CC antibodies to Factor VIIIc, and in diagnostic assays for the presence of
XX CC Factor VIIIc subunits in physiological fluids, for example blood or

```

```

CC serum.
XX SQ Sequence 211 AA:

Query Match      86.7%; Score 65; DB 21; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.00055;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVYLRKHQR 14
   |||||
DB 115 sqnpvylkrhqr 126

RESULT 12
AAB07205
XX ID AAB07205 standard; Protein; 1283 AA.
XX AC AAB07205;
XX DT 11-OCT-2000 (first entry)
XX DE Human Factor VIIIc protein sequence.
XX KW Factor VIII C domain; human; blood clotting; haemophilia.
XX OS Homo sapiens.
XX PA (CHIR ) CHIRON CORP.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Kuo G, Rasmussen ME, Maslarsz FR, Valenzuela P, Truett M;
XX PI Favalloro J;
XX DR WPI; 2000-367967/32.
XX DR N-PSDB; AAA58443.
XX PT DNA sequence of portion of human Factor VIIIc for treating and
XX PT preventing the symptoms of haemophilia -
XX PS Disclosure; Page 31-35; 39pp; English.
XX CC The present sequence is the protein sequence for human Factor VIIIc,
XX CC which was obtained by sequencing a genomic DNA library. Factor VIIIc is a
XX CC plasma protein involved in blood coagulation, and is absent or defective
XX CC in haemophilia A. The Factor VIIIc protein can, therefore, be used to
XX CC treat haemophilia, as well as in the production of monoclonal antibodies
XX CC to Factor VIIIc, and in diagnostic assays for the presence of Factor
XX CC VIIIc subunits in physiological fluids, for example blood or serum.
XX SQ Sequence 1283 AA:

Query Match      86.7%; Score 65; DB 21; Length 1283;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVYLRKHQR 14
   |||||
DB 952 sqnpvylkrhqr 963

```

```

PR 01-AUG-1986; 86US-0893375.
XX
PA (BIOI ) BIOGEN NV (PASE/).
XX
PI Pasek MP;
XX
DR WPI; 1988-049866/07.
XX
DR N-PSDB; AAN80444.
XX
PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA
XX encoding maturation polypeptide, useful for high yield transformation.
XX
PS Claim 3; Page 51-52-53-54; 97p; English.
XX
CC A major part of the sequence encoding the maturation polypeptide of
CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563. The QD deletion
CC retains approximately 90 amino acids of the maturation polypeptide
CC (four amino acids at the N-terminal end and 86 amino acids at
CC the C-terminal end). The full length factor VIII:C cDNA has two
CC changes with respect to the published sequence (EPO application 160457):
CC CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880
CC (Phe to Leu)
CC
CC The product is produced in approx. 20 times higher
CC yields than previous recombinant produced factor VIII:C and are more
CC easily purified. The peptide is used for treating haemophilia A, both
CC acute and prolonged bleeding.
CC See also AAN80446 and AAN80447.
XX
XX
SQ Sequence 1516 AA;
XX
XX
Query Match 86.7%; Score 65; DB 9; Length 1516;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SONPPYIKRHQR 14
| | | | | | | | | |
Db 821 sqnppylkrhqr 832
XX
RESULT 15
AAM18670
ID AAM18670 standard; Protein; 1661 AA.
XX
AC AAM18670;
XX
DT 10-AUG-1997 (first entry)
XX
DE Factor VIII-dB695-HCII.
XX
KW Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
KW blood clotting; procoagulant; anticoagulant; antithrombotic;
KW haemophilia; gene therapy.
XX
XX Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FH Domain 731..760
FT /label=HCII
FT /note="heparin cofactor II acidic domain"
XX
PN MO9718315-A1.
XX
PD 22-MAY-1997.
XX
PE 13-NOV-1996; 96WO-EP04977.
XX
PR 13-NOV-1995; 95US-0558107.
XX
PA (IMMO ) IMMUNO AG.
XX
PI Voorberg JJ;
XX

```

DR WPI: 1997-289291/26.
DR N-PSDB: AAT69811.

XX Hybrid Factor VIII with modified activity, comprises region from
PT donor anticoagulant or antithrombotic protein - useful for treatment
PT of coagulation disorders

XX Claim 11; Page 52-60; 96pp; English.

PS
XX Factor VIII-dB695-HCII (AAM18670) is a hybrid protein in which amino
CC acids 712-736 of Factor-dB695 (Factor VIII del868-1562) B-domain
CC are replaced by amino acids 51-80 from the acidic region (and
CC potential thrombin-binding site) of human heparin cofactor II
CC (HCII). It is the expression product of Factor VIII-dB695-HCII
CC DNA (AAT69811) in plasmid pCLB-dB695-HCII. The hybrid protein, which
CC can be expressed using gene therapy techniques, has increased
CC procoagulant activity owing to the HCII acidic region, and can be
CC used to treat blood coagulation disorders such as haemophilia A.
XX
SQ Sequence 1661 AA;

Query Match .86.7%; Score 65; DB 18; Length 1661;

Best Local Similarity 100.0%; Pred. No. 0.0051;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLKRHR 14
|||||

DB 966 sqnppvlkrhgr 977

Search completed: January 17, 2002, 12:36:28
Job time: 33 sec

Tue Jan 22 15:12:52 2002

us-09-740-211-15.ra1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:35:55 ; Search time 12.5 seconds
(Without alignments)
25.204 Million cell updates/sec

Title: US-09-740-211-15

Sequence: 1 SFSQNPVLKRHR 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued_Patents_AA.*

1: /cgn2-6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2-6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2-6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2-6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2-6/ptodata/2/1aa/PCrus.COMB.pep.*
6: /cgn2-6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	75	100.0	14	US-08-634-001C-1
2	75	100.0	14	US-09-470-618-15
3	65	86.7	1661	US-08-882-083-2
4	65	86.7	1661	US-08-558-107-2
5	65	86.7	1661	US-09-243-539-2
6	65	86.7	2332	US-07-864-004B-4
7	65	86.7	2332	US-08-251-937A-4
8	65	86.7	2332	US-08-212-133A-2
9	65	86.7	2332	US-08-276-594A-2
10	65	86.7	2332	US-08-474-503-2
11	65	86.7	2332	US-08-670-707A-2
12	65	86.7	2332	US-09-037-601-2
13	65	86.7	2332	US-09-324-867-3
14	65	86.7	2332	US-09-324-867-4
15	65	86.7	2332	US-09-324-867-5
16	65	86.7	2332	US-09-324-867-6
17	65	86.7	2332	US-09-324-867-7
18	65	86.7	2332	US-09-324-867-8
19	65	86.7	2332	US-09-324-867-9
20	65	86.7	2332	US-09-324-867-10
21	65	86.7	2332	US-09-324-867-11
22	65	86.7	2332	US-09-324-867-12
23	65	86.7	2332	US-09-324-867-13
24	65	86.7	2332	US-09-324-867-14
25	65	86.7	2332	US-09-324-867-15
26	65	86.7	2332	US-09-324-867-16
27	65	86.7	2332	US-09-324-867-17

28	55	73.3	34	US-08-441-943-30	Sequence 30, Appl
29	54	72.0	4	US-09-324-867-2	Sequence 2, Appl
30	50	66.7	9	US-08-441-943-34	Sequence 34, Appl
31	47	62.7	868	US-07-864-004B-6	Sequence 6, Appl
32	47	62.7	868	US-08-251-937A-6	Sequence 6, Appl
33	47	62.7	868	US-08-212-133A-3	Sequence 6, Appl
34	47	62.7	1090	US-09-324-867-5	Sequence 6, Appl
35	47	62.7	2115	US-09-324-867-5	Sequence 6, Appl
36	47	62.7	2133	US-08-670-707A-37	Sequence 37, Appl
37	47	62.7	2133	US-09-037-601-37	Sequence 37, Appl
38	38	50.7	364	US-08-318-831-6	Sequence 6, Appl
39	38	50.7	1333	US-09-356-952-2	Sequence 2, Appl
40	37	49.3	158	US-09-374-135-2	Sequence 2, Appl
41	36	48.0	274	US-08-248-466B-10	Sequence 10, Appl
42	36	48.0	351	US-08-248-466B-12	Sequence 12, Appl
43	36	48.0	392	US-09-416-050A-2	Sequence 2, Appl
44	36	48.0	392	US-09-664-800-2	Sequence 2, Appl
45	36	48.0	392	US-09-665-309-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-634-001C-1

Sequence 1, Application US/08634001C

Patent No. 5952198

GENERAL INFORMATION:

APPLICANT: Chan, Sham-Yuen

TITLE OF INVENTION: Production of Recombinant Factor VIII

TITLE OF INVENTION: In the Presence of Liposome-Like

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bayer Corporation

STREET: 800 Dwight Way

CITY: Berkeley

STATE: California

COUNTRY: USA

ZIP: 94701-1986

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB Storage

COMPUTER: IBM

OPERATING SYSTEM: DOS

SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/634,001C

FILING DATE: 4/35

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/434,900

FILING DATE: May 4, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Giblin, James A.

REGISTRATION NUMBER: 25772

REFERENCE/DOCKET NUMBER: MSB-7226CIP

TELEPHONE: (510)705-7910

TELEFAX: (510)705-7904

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 14

TYPE: amino acid

STRANDEDNESS: single strand

TOPOLOGY: linear

MOLECULE TYPE: peptide

DESCRIPTION: peptide

US-08-634-001C-1

W0 9169121

Query Match 100.0%; Score 75; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSQNPVLRKHOR 14
Db 1 SFSQNPVLRKHOR 14

RESULT 2
US-09-470-618-15

Sequence 15, Application US/09470618
Patent No. 6200560
GENERAL INFORMATION:
APPLICANT: Couto, Linda B.
APPLICANT: Colosi, Peter C.
TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
FILE REFERENCE: Avigen-04082
CURRENT APPLICATION NUMBER: US/09/470,618
EARLIER FILING DATE: 1999-12-22
EARLIER APPLICATION NUMBER: 09/364,862
EARLIER FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: 60/125,974
EARLIER FILING DATE: 1999-03-24
EARLIER APPLICATION NUMBER: 60/104,994
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 15
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-15

Query Match 100.0%; Score 75; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSQNPVLRKHOR 14
Db 1 SFSQNPVLRKHOR 14

RESULT 3
US-08-882-083-2

Sequence 2, Application US/08882083
Patent No. 5869292
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,083
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:

NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-882-083-2

Query Match 86.7%; Score 65; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVLRKHOR 14
Db 966 SQNPVLRKHOR 977

RESULT 4
US-08-558-107-2

Sequence 2, Application US/08558107
Patent No. 5910481
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,107
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-107-2

Query Match 86.7%; Score 65; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVLRKHOR 14
Db 966 SQNPVLRKHOR 977

RESULT 5
US-09-243-539-2
Sequence 2, Application US/09243539
Patent No. 6130203
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,539
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-243-539-2

Query Match 86.7%; Score 65; DB 4; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVLRHQR 14
|||||
Db 966 SQNPVLRHQR 977

RESULT 6
US-07-864-004B-4
Sequence 4, Application US/07864004B
Patent No. 5364771
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: EM0106
APPLICATION NUMBER: US/07/864,004B
FILING DATE: 07 APRIL 1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EM0106
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ. ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HIPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver cDNA sequence
US-07-864-004B-4

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVLRHQR 14
|||||
Db 1637 SQNPVLRHQR 1648

RESULT 7
US-08-251-937A-4
Sequence 4, Application US/08251937A
Patent No. 5583209
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/251,937A
FILING DATE: 31-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EM0106DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6367

TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver CDNA sequence
US-08-251-937A-4

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLRKHQR 14
|||||
Db 1637 SONPPVLRKHQR 1648

RESULT 8
US-08-212-133A-2
Sequence 2, Application US/08212133A
Patent No. 5663060
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,133A
FILING DATE: March 11, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU/76677
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien

TISSUE TYPE: Liver CDNA sequence
US-08-212-133A-2

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLRKHQR 14
|||||
Db 1637 SONPPVLRKHQR 1648

RESULT 9
US-08-276-594A-2
Sequence 2, Application US/08276594A
Patent No. 5693499
GENERAL INFORMATION:
APPLICANT: YONEMURA, Hiroshi
APPLICANT: TAJIMA, Yoshitaka
APPLICANT: SUGAWARA, Keishin
APPLICANT: MASUDA, Kenichi
TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
FACTOR VIII PROTEIN COMPLEX
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,594A
FILING DATE: 18-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/950,191
FILING DATE: 24-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 243262/1991
FILING DATE: 24-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 74129/195/NOA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-276-594A-2

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLRKHQR 14
|||||
Db 1637 SONPPVLRKHQR 1648

RESULT 10
US-08-474-503-2
Sequence 2, Application US/08474503
Patent No. 5744446
GENERAL INFORMATION:
APPLICANT: Emory University
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,503
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EMU106CIP(3)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6500
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver cdna sequence
US-08-474-503-2

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPVYKRRHR 14
|||||

Db 1637 SONPVYKRRHR 1648

RESULT 11
US-08-670-707A-2
Sequence 2, Application US/08670707A
Patent No. 5859204
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,707A
FILING DATE: 26-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Greenlee, Lorraine L.
REGISTRATION NUMBER: 27,894
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Liver
US-08-670-707A-2

Query Match 86.7%; Score 65; DB 2; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPVYKRRHR 14
|||||

Db 1637 SONPVYKRRHR 1648

RESULT 12
US-09-037-601-2
Sequence 2, Application US/09037601
Patent No. 6180371
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Liver
US-09-037-601-2

Query Match 86.7%; Score 65; DB 4; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLKRHR 14
|||||

DB 1637 SONPPVLKRHR 1648

RESULT 13
US-09-324-867-3
Sequence 3, Application US/09324867A
Patent No. 6251632
GENERAL INFORMATION:
APPLICANT: Lillietap, David
APPLICANT: Camerton, Cherle
APPLICANT: No. 6251632ley, Colleen
APPLICANT: Horrocks, L. Suzanne Hoyle
APPLICANT: Hough, Christine
TITLE OF INVENTION: Carline Factor VIII Gene, Protein and Methods of Use
FILE REFERENCE: 1669.0010002/JAG/BJD
CURRENT APPLICATION NUMBER: US/09/324,867A
CURRENT FILING DATE: 1999-06-03
EARLIER APPLICATION NUMBER: 09/035,141
EARLIER FILING DATE: 1998-03-059
EARLIER APPLICATION NUMBER: 60/039,953
EARLIER FILING DATE: 1997-03-06
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 2332
TYPE: PRT
ORGANISM: Homo sapiens
US-09-324-867-3

Query Match 86.7%; Score 65; DB 4; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SONPPVLKRHR 14
|||||

DB 1637 SONPPVLKRHR 1648

RESULT 14
PCT-US93-03275-4
Sequence 4, Application PC/TUS9303275
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03275
FILING DATE: 19930407
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU 106PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver CDNA sequence
PCT-US93-03275-4

Query Match 86.7%; Score 65; DB 5; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLKRHR 14
|||||

DB 1637 SONPPVLKRHR 1648

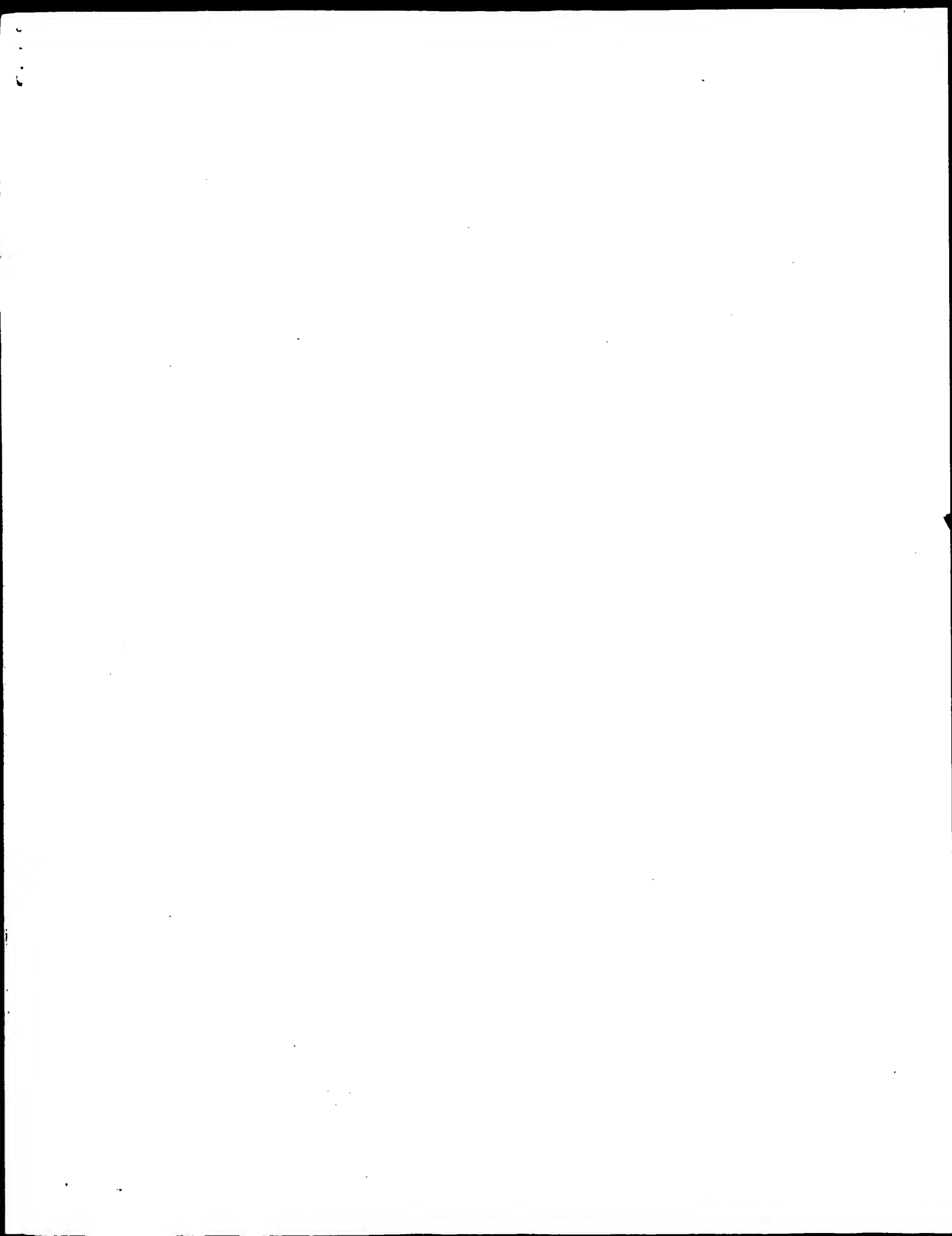
RESULT 15
PCT-US94-13200-2
Sequence 2, Application PC/TUS9413200
GENERAL INFORMATION:
APPLICANT: Emory University
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta

STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/13200
FILING DATE: 15-NOV-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU106CIP(2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver cdna sequence
PCT-US94-13200-2

Query Match 86.7%; Score 65; DB 5; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SGNPPVLKRHR 14
Db 1637 SGNPPVLKRHR 1648

Search completed: January 17, 2002, 12:36:47
Job time: 52 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:36:00 ; Search time 10.27 Seconds
(without alignments)
49.981 Million cell updates/sec

Title: US-09-740-211-15

Sequence: 1 SEQONPPVLRKRROR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	65	86.7	2351 1	FA8_HUMAN
2	56	74.7	2319 1	FA8_MOUSE
3	47	62.7	2133 1	FA8_PIG
4	41	54.7	2415 1	SPCA_DROME
5	38	50.7	199 1	R13A_YEAST
6	38	50.7	199 1	R13B_YEAST
7	38	50.7	304 1	GBD1_WHEAT
8	38	50.7	436 1	Y177_TREPA
9	38	50.7	1333 1	SOS1_HUMAN
10	37	48.3	356 1	G17A_WHEAT
11	36	48.0	94 1	P16B_MOUSE
12	36	48.0	311 1	PROC_NEUCR
13	36	48.0	366 1	HTPX_AERPE
14	36	48.0	394 1	BMP2_MOUSE
15	36	48.0	561 1	YGG4_YEAST
16	36	48.0	737 1	OPT1_DROME
17	36	48.0	873 1	CPHA_SYNY3
18	36	48.0	1522 1	MRP3_RAT
19	36	48.0	2128 1	SPCB_MOUSE
20	35.5	47.3	1833 1	ZEP2_HUMAN
21	35	46.7	292 1	HTPX_PYRO
22	35	46.7	314 1	NODD_AZOC
23	35	46.7	516 1	Y067_MYCE
24	35	46.7	529 1	DNB2_ADE02
25	35	46.7	529 1	DNB2_ADE05
26	35	46.7	533 1	LCP2_MOUSE
27	35	46.7	549 1	CEP1_YEAST
28	35	46.7	614 1	ZF28_MOUSE
29	35	46.7	656 1	YAN9_SCHPO
30	35	46.7	658 1	UVRD_RICPR
31	35	46.7	743 1	MYBB_XENLA
32	35	46.7	758 1	PKA1_YEAST
33	35	46.7	933 1	KGP3_DROME

34	35	46.7	1014 1	UVRA_SNRCO	O92507 streptomyc
35	35	46.7	1184 1	CHSD_EMENT	P78611 emericella
36	35	46.7	1319 1	SOS1_MOUSE	O62245 mus musculi
37	35	46.7	1716 1	RPA1_RAT	O54689 rattus norv
38	35	46.7	4486 1	DYH9_HUMAN	O94959 homo sapien
39	34.5	46.0	580 1	ZF35_MOUSE	P15620 mus musculi
40	34.5	46.0	711 1	2175_HUMAN	O94473 homo sapien
41	34	45.3	105 1	RL24_THEMA	P38513 thermotoga
42	34	45.3	170 1	VIP_MOUSE	P32648 mus musculi
43	34	45.3	215 1	Y042_METUA	O60347 methanococ
44	34	45.3	309 1	YL11_CAEEL	Q11098 caenorhabd
45	34	45.3	322 1	ASTE_ECOLI	P76215 escherichia

ALIGNMENTS

RESULT 1	FA8_HUMAN	STANDARD:	PRT: 2351 AA.
ID	FA8_HUMAN		
AC	P00451		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT)		
DE	(ANTHEMOPHILIC FACTOR) (AHF).		
GN	F8 OR F8C.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86081164; PubMed=3935400;		
RA	Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,		
RA	Hartog K., Kuo C.H., Maslarz F.R., Merryweather J.P., Najarian R.,		
RA	Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,		
RA	Urdea M.S., Valenzuela P., Dahl H.-H.M., Favaloro J., Hansen J.,		
RA	Nordfang O., Ezban M.		
RT	"Characterization of the polypeptide composition of human factor		
RT	VIII:C and the nucleotide sequence and expression of the human kidney		
RT	CDNA.";		
RL	DNA 4:333-349(1985).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85061548; PubMed=6438526;		
RA	Wood W.L., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,		
RA	Key B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,		
RA	Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;		
RT	"Expression of active human factor VIII from recombinant DNA clones.";		
RT	Nature 312:330-337(1984).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85061550; PubMed=6438528;		
RA	Toole J.D., Knopf J.L., Wozney J.M., Soltzman L.A., Buecker J.L.,		
RA	Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,		
RA	Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,		
RA	Hewick R.M.;		
RT	"Molecular cloning of a cDNA encoding human antithrombophilic factor.";		
RT	Nature 312:342-347(1984).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93265012; PubMed=1303178;		
RA	Gitschier J., Wood W.L.;		
RT	"Sequence of the exon-containing regions of the human factor VIII		
RT	gene.";		
RL	Hum. Mol. Genet. 1:199-200(1992).		
RN	[5]		
RP	SEQUENCE OF 2064-2070 FROM N.A.		
RA	de Water N.S., Williams R., Browett P.J.;		
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.		
RN	[6]		
RP	SULFATION OF 1699.		

RX MEDLINE=91093266; PubMed=1898735;
 RA Leyte A., van Schijndel H.B., Niehs C., Hutner W.B., Verbeet M.P.,
 RA Mertens K., van Mourik J.A.;
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor";
 RL J. Biol. Chem. 266:740-746(1991).
 RN [17]
 RP SOLFATION.
 RX MEDLINE=92207952; PubMed=1554716;
 RA Pittman D.D., Wang J.H., Kaufman R.J.;
 RT "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII";
 RL Biochemistry 31:3315-3325(1992).
 RN [18]
 RP STRUCTURE BY NMR OF 2322-2343.
 RX MEDLINE=95200924; PubMed=7893714;
 RA Gilbert G.E., Baleja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy";
 RL Biochemistry 34:3022-3031(1995).
 RN [19]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=91221499; PubMed=1902642;
 RA Gitschier J.;
 RT "The molecular basis of hemophilia A";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [110]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A";
 RL Blood 73:1-12(1989).
 RN [111]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=95245332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A";
 RL Hum. Mutat. 5:1-22(1995).
 RN [112]
 RP VARIANT GLN-2326.
 RX MEDLINE=86235434; PubMed=3012775;
 RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophilia";
 RL Science 233:1415-1416(1986).
 RN [113]
 RP VARIANT PRO-2135.
 RX MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [114]
 RP VARIANT GLN-2228.
 RX MEDLINE=88191889; PubMed=2833855;
 RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides";
 RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [115]
 RP VARIANT GLY-291.
 RX MEDLINE=88220354; PubMed=2835904;
 RA Youssoufian H., Wong C., Aronis S., Platakoukis H., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
 RT in exon 7 of the factor VIII gene";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [116]
 RP VARIANT CYS-1708.
 RX MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;
 RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A";
 RL Blood 73:2117-2122(1989).
 RN [117]
 RP VARIANT CYS-391.
 RX MEDLINE=90001543; PubMed=2506948;
 RA Shima M., Ware J., Yoshioka A., Fukui H., Fuleher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule";
 RL Blood 74:1612-1617(1989).
 RN [118]
 RP VARIANT LEU-189.
 RX MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII C gene
 RT resulting in moderately severe hemophilia A";
 RL Blood 74:2688-2691(1989).
 RN [119]
 RP VARIANT LEU-2326.
 RX MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene";
 RL Hum. Genet. 81:335-338(1989).
 RN [120]
 RP VARIANT HIS-391.
 RX MEDLINE=89264602; PubMed=2498882;
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372-->histidine)";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [121]
 RP VARIANT CYS-1708.
 RX MEDLINE=90105723; PubMed=2104766;
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A";
 RL Blood 75:384-389(1990).
 RN [122]
 RP VARIANTS GLN-2228 AND LEU-2326.
 RX MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent";
 RL Blood 75:662-670(1990).
 RN [123]
 RP VARIANT CYS-391.
 RX MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CRM+ haemophilia A due to a missense mutation (372-->Cys) at the
 RT internal heavy chain thrombin cleavage site";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [124]
 RP VARIANTS PHE-1699 AND CYS-1708.
 RX MEDLINE=90152691; PubMed=2105906;
 RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
 RA Kazazian H.H., Antonarakis S.E.;
 RT "Characterization of mutations in the factor VIII gene by direct
 RT sequencing of amplified genomic DNA";
 RL Genomics 6:65-71(1990).
 RN [125]
 RP VARIANTS CYS-1728 AND ASP-1941.
 RX MEDLINE=90169988; PubMed=2106480;
 RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
 RA Kazazian H.H.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene";

Query Match 86.7%; Score 65; DB 1; Length 2351;
 Best Local Similarity 100.0%; Pred. No. 0.0035;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVLRHR 14
 |||||||||
 Db 1656 SQNPVLRHR 1667

RESULT 2
 ID FA8_MOUSE STANDARD; PRT: 2319 AA.

AC 006194;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT).
 GN CF8 OR F8C.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RA MEDLINE=93300511; PubMed=8314577;
 RA Elder B., Laskich D., Gitschler J.;
 RT "Sequence of the murine factor VIII cDNA";
 RL Genomics 16:374-379(1993).
 CC -1 FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
 AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
 CC ACTIVATED FORM, FACTOR Xa.
 CC -1 SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1 TISSUE SPECIFICITY: FOUND IN MOST TISSUES.
 CC -1 SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF
 2 PLASTOCYANIN-LIKE REPEATS.
 CC -1 SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1 SIMILARITY: STRONG, TO COAGULATION FACTOR V.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use, by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: L05573; AAA37385.1; -
 DR PIR: A47004; A47004.
 DR HSSP: P00451; ICFG.
 DR MGD: MGI:88363; F8.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR000421; FA58_C.
 DR Pfam: PF00394; Cu-oxidase: 3.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR SMART: SM00231; FA58C; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 KW Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
 KW Signal; Glycoprotein; Sulfation.
 KM SIGNAL: Glycoprotein; Sulfation.
 FT CHAIN 1
 FT CHAIN 19
 FT CHAIN 20 2319 COAGULATION FACTOR VIII.
 FT DOMAIN 20 349 F5/8 TYPE A 1.
 FT DOMAIN 20 199 PLASTOCYANIN-LIKE 1.
 FT DOMAIN 207 349 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 399 730 F5/8 TYPE A 2.
 FT DOMAIN 399 573 PLASTOCYANIN-LIKE 3.
 FT DOMAIN 583 730 PLASTOCYANIN-LIKE 4.
 FT DOMAIN 760 1640 B.
 FT DOMAIN 1683 2008 F5/8 TYPE A 3.
 FT DOMAIN 1683 1845 PLASTOCYANIN-LIKE 5.

FT DOMAIN 1855 2008 PLASTOCYANIN-LIKE 6.
 FT DOMAIN 2008 2156 F5/8 TYPE C 1.
 FT DOMAIN 2161 2313 F5/8 TYPE C 2.
 FT SITE 391 392 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 759 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1678 1679 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1324 1325 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 FT SITE 1640 1641 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 FT MOD_RES 367 367 SULFATION (BY SIMILARITY).
 FT MOD_RES 737 737 SULFATION (BY SIMILARITY).
 FT MOD_RES 738 738 SULFATION (BY SIMILARITY).
 FT MOD_RES 742 742 SULFATION (BY SIMILARITY).
 FT MOD_RES 1669 1669 SULFATION (REQUIRED FOR VWF BINDING) (BY SIMILARITY).
 FT MOD_RES 1667 1667 SULFATION (BY SIMILARITY).
 FT MOD_RES 173 199 PROBABLE.
 FT DISULFID 547 573 PROBABLE.
 FT DISULFID 1819 1845 PROBABLE.
 FT DISULFID 2008 2156 BY SIMILARITY.
 FT DISULFID 2161 2313 BY SIMILARITY.
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 958 958 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1026 1026 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1044 1044 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1087 1087 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1136 1136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1161 1161 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1192 1192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1255 1255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1274 1274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1302 1302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1316 1316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1340 1340 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1797 1797 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2105 2105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 2319 AA; 266148 MW; FD054DBD051DB2A01 CRC64;

Query Match 74.7%; Score 56; DB 1; Length 2319;
 Best Local Similarity 91.7%; Pred. No. 0.13;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 SQNPVLRHR 14
 |||||||||
 Db 1629 SQNPVLRHR 1640

RESULT 3
 ID FA8_PIG STANDARD; PRT: 2133 AA.

AC P12263; Q95243;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT).
 GN CF8.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.

Query Match	Best Local Similarity	Score 47;	DB 1;	Length 2133;
Matches 8;	Conservative	88.9%;	Pred. No. 4.3;	Mismatches 0;
Indels	0;	Gaps	0;	
Db	1441 PVLKRROR 1449			
Qy	6 PVLKRROR 14			
RESULT 4				
SPCA_DROME	STANDARD;	PRT;	2415 AA.	
ID	SPCA_DROME			
AC	P13395; Q9M085; Q26340;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	SPECTRIN ALPHA CHAIN.			
GN	ALPHA-SPEC OR SPEC-A OR CG1977.			
OS	Drosophila melanogaster (Fruit fly).			
OS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha			
OC	Phylogridea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90037215; PubMed=2806524;			
RA	Dubreuil R.R., Byers T.J., Sillman A.L., Bar-Zvi D.,			
RA	Goldstein L.S.B., Branton D.,			
RT	"The complete sequence of Drosophila alpha-spectrin: conservation of			
RT	structural domains between alpha-spectrins and alpha-actinin.";			
RU	J. Cell Biol. 109:2197-2205(1989).			
RU	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BERKELEY;			
RC	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	Georgie R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Franko C., Baldwin D.,			
RA	Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,			
RA	Botkova D., Botchan R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
FT	CARBOHD 259	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 601	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 929	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 985	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1025	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1111	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1181	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1208	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1245	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1265	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1335	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1408	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1611	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CARBOHD 1919	N-LINKED (GLCNAC. . .)	(POTENTIAL).	
FT	CONFLICT 713	N-> M (IN REF. 2).	(POTENTIAL).	
FT	CONFLICT 734	I-> T (IN REF. 2).		
FT	CONFLICT 792	G-> Q (IN REF. 2).		
FT	CONFLICT 1133	E-> F (IN REF. 2).		
FT	CONFLICT 1191	I-> L (IN REF. 2).		
FT	CONFLICT 1209	R-> F (IN REF. 2).		
FT	CONFLICT 1437	C-> G (IN REF. 2).		
FT	CONFLICT 1436	F-> R (IN REF. 2).		
FT	CONFLICT 1539	F-> N (IN REF. 2).		
FT	CONFLICT 1546	Q-> N (IN REF. 2).		
SO	SEQUENCE 2133 AA; 239304 MW; 152BBA8997F570DA CRC64;			

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 1-150 AND 2192-2415 FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=94103334; PubMed=8276898;
 RA Lee J.R., Coyne R.S., Dubreuil R.R., Goldstein L.S.B., Branton D.;
 RT "Cell shape and interaction defects in alpha-spectrin mutants of
 RT *Drosophila melanogaster*.";
 RL J. Cell Biol. 123:1797-1809(1993).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=88059242; PubMed=3680372;
 RA Dubreuil R., Byers T.J., Branton D., Goldstein L.S.B., Kiehart D.P.;
 RT "Drosophila spectrin. I. Characterization of the purified protein.";
 RL J. Cell Biol. 105:2095-2102(1987).
 RN [5]
 RP EMBRYONIC LOCALIZATION.
 RX MEDLINE=89234159; PubMed=2497103;
 RA Pesecreta T.C., Byers T.J., Dubreuil R., Kiehart D.P., Branton D.;
 RT "Drosophila spectrin. I. The membrane skeleton during embryogenesis.";
 RL J. Cell Biol. 108:1697-1709(1989).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1391-1497.
 RX MEDLINE=94093440; PubMed=8266097;
 RA Yan Y., Winograd E., Viel A., Cronin T., Harrison S.C., Branton D.;
 RT "Crystal structure of the repetitive segments of spectrin.";
 RL Science 262:2027-2030(1993).
 CC -1- FUNCTION: SPECTRIN IS THE MAJOR CONSTITUENT OF THE CYTOSKELETAL
 CC NETWORK UNDERLYING THE ERYTHROCYTE PLASMA MEMBRANE. IT ASSOCIATES
 CC WITH BAND 4.1 AND ACTIN TO FORM THE CYTOSKELETAL SUPERSTRUCTURE OF
 CC THE ERYTHROCYTE PLASMA MEMBRANE. ESSENTIAL FOR LARVAL SURVIVAL AND
 CC DEVELOPMENT. STABILIZES CELL TO CELL INTERACTIONS THAT ARE
 CC CRITICAL FOR THE MAINTENANCE OF CELL SHAPE AND SUBCELLULAR
 CC ORGANIZATION WITHIN EMBRYONIC TISSUES.
 CC -1- SUBUNIT: NATIVE SPECTRIN MOLECULE IS A Tetramer composed of two
 CC ANTIPARALLEL HETERODIMERS JOINED HEAD TO HEAD SO THAT EACH END OF
 CC THE NATIVE MOLECULE INCLUDES THE C-TERMINUS OF THE ALPHA SUBUNIT
 CC AND THE N-TERMINUS OF THE BETA SUBUNIT. INTERACTS WITH CALMODULIN
 CC IN A CALCIUM-DEPENDENT MANNER AND INTERACTS WITH F-ACTIN.
 CC -1- SUBCELLULAR LOCATION: NEAR THE INNER SURFACE OF THE PLASMA
 CC MEMBRANE OF NEARLY ALL CELLS.
 CC -1- TISSUE SPECIFICITY: A SUBSTANTIAL POOL OF MATERNAL PROTEIN IN THE
 CC EGG UNDERGOES DYNAMIC CHANGES IN DISTRIBUTION EARLY IN
 CC EMBRYOGENESIS. IN GASTRULATED EMBRYO, THE HIGHEST LEVEL OF PROTEIN
 CC IS FOUND IN THE RESPIRATORY TRACT CELLS AND THE LOWEST IN PARTS OF
 CC THE FORMING GUT.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
 CC -1- SIMILARITY: BELONGS TO THE SPECTRIN FAMILY.

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 22 SPECTRIN REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M26400; AAA28907.1; -
 CC EMBL: AE003472; AAF7569.1; -
 CC EMBL: S67762; AAB29441.2; -
 CC EMBL: S67765; AAB29442.1; -
 CC PIR: A3733; A37333.
 CC PDB: 2SPC; 3I-MAY-94.
 CC FLYBase: FBgn0003470; Alpha-Spec.
 CC InterPro: IPR002048; EF-hand.
 CC InterPro: IPR001452; SH3.
 CC InterPro: IPR002017; Spectrin.
 CC Pfam: PF00036; efhand; 2.
 CC Pfam: PF00018; SH3; 1.
 CC Pfam: PF00435; spectrin; 22.
 CC PRINTS: PR00452; SH3DOMAIN.
 CC SMART: SM00054; Efh; 2.
 CC SMART: SM00326; SH3; 1.
 CC SMART: SM00150; SPEC; 20.
 CC PROSITE: PS00018; EF_HAND; 2.
 CC PROSITE: PS50002; SH3; 1.
 CC KW Cytoskeleton; Membrane; Erythrocyte; Repeat; Actin-binding;
 CC KW Capping protein; Calcium-binding; SH3 domain; 3D-structure;
 CC KW Cell shape; Calmodulin-binding.
 CC FT REPEAT 18 122 SPECTRIN 1.
 CC FT REPEAT 123 228 SPECTRIN 2.
 CC FT REPEAT 229 334 SPECTRIN 3.
 CC FT REPEAT 335 440 SPECTRIN 4.
 CC FT REPEAT 441 546 SPECTRIN 5.
 CC FT REPEAT 547 651 SPECTRIN 6.
 CC FT REPEAT 652 757 SPECTRIN 7.
 CC FT REPEAT 758 863 SPECTRIN 8.
 CC FT REPEAT 864 969 SPECTRIN 9.
 CC FT REPEAT 970 1043 SPECTRIN 10.
 CC FT REPEAT 1044 1151 SPECTRIN 11.
 CC FT REPEAT 1152 1257 SPECTRIN 12.
 CC FT REPEAT 1258 1363 SPECTRIN 13.
 CC FT REPEAT 1364 1469 SPECTRIN 14.
 CC FT REPEAT 1470 1576 SPECTRIN 15.
 CC FT REPEAT 1577 1682 SPECTRIN 16.
 CC FT REPEAT 1683 1788 SPECTRIN 17.
 CC FT REPEAT 1789 1894 SPECTRIN 18.
 CC FT REPEAT 1895 2001 SPECTRIN 19.
 CC FT REPEAT 2002 2115 SPECTRIN 20.
 CC FT REPEAT 2116 2229 SPECTRIN 21.
 CC FT REPEAT 2230 2335 SPECTRIN 22.
 CC FT CA_BIND 2278 2289 EF_HAND 1 (POTENTIAL).
 CC FT CA_BIND 2321 2332 EF_HAND 2 (POTENTIAL).
 CC FT DOMAIN 970 1029 SH3.
 CC FT CONFLICT 110 110 Q -> D (IN REF. 3).
 CC FT CONFLICT 1668 1668 Q -> R (IN REF. 1).
 CC SO SPOUNCE 2415 AA; 278301 MW; F1F72FB990E0A7 CRC64;

Query Match 54.7%; Score 41; DB 1; Length 2415;
 Best local 58.3%; Pred. No. 55;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 SFSQNPVLKRH 12
 Db 1995 SHAQSPAILKRH 2006
 RESULT 5

```

R13A_YEAST
ID R13A_YEAST STANDARD; PRT; 199 AA.
AC 012690;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S RIBOSOMAL PROTEIN L13-A.
GN RPL13A OR YD1082W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Wambolt R., Wedler H., Wedler E., Scharfe M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L13 IN YEAST.
CC -1- SIMILARITY: BELONGS TO THE L13 FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 274130; CAA98648.1; -.
CC DR SGD; S0002240; RPL13A.
CC DR InterPro; IPR001380; Ribosomal_L13E.
CC DR Pfam; PF01294; Ribosomal_L13E; 1.
CC DR ProDom; PD004443; Ribosomal_L13E; 1.
CC DR PROSITE; PS01104; RIBOSOMAL_L13E; 1.
CC KW Ribosomal protein; Multigene family.
CC SEQUENCE 199 AA; 22554 MW; EAVCCSFIF3C0335A CRC64;
SQ
Query Match 50.7%; Score 38; DB 1; Length 199;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHQR 14
Db 2 AISKNPILKNHFR 15
RESULT 6
ID R13B_YEAST STANDARD; PRT; 199 AA.
AC P40212;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S RIBOSOMAL PROTEIN L13-B.
GN RPL13B OR RPL13 OR YMR142C OR YM9375.11C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-6288C / AB972;
RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L13 IN YEAST.
CC -1- SIMILARITY: BELONGS TO THE L13 FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 247071; CAA87356.1; -.
CC DR SGD; S0004750; RPL13B.
CC DR InterPro; IPR001380; Ribosomal_L13E.
CC DR Pfam; PF01294; Ribosomal_L13E; 1.
CC DR ProDom; PD004443; Ribosomal_L13E; 1.
CC DR PROSITE; PS01104; RIBOSOMAL_L13E; 1.
CC KW Ribosomal protein; Multigene family.
CC SEQUENCE 199 AA; 22525 MW; BEA3B423EE76EF23 CRC64;
SQ
Query Match 50.7%; Score 38; DB 1; Length 199;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHQR 14
Db 2 AISKNPILKNHFR 15
RESULT 7
ID GDB1_WHEAT STANDARD; PRT; 304 AA.
AC P04729;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GAMMA-GLIADIN B-I PRECURSOR.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-85234522; PubMed-2989281;
RA Okita T.W., Cheesbrough V., Reeves C.D.;
RT "Evolution and heterogeneity of the alpha-/beta-type and gamma-type
RT gliadin DNA sequences";
RL J. Biol. Chem. 260:8203-8213(1985).
CC -1- FUNCTION: GLIADIN IS THE MAJOR SEED STORAGE PROTEIN IN WHEAT.
CC -1- MISCELLANEOUS: THE GAMMA-GLIADINS CAN BE DIVIDED INTO 3 HOMOLOGY
CC CLASSES. SEQUENCE DIVERGENCE BETWEEN THE CLASSES IS DUE TO
CC SINGLE-BASE SUBSTITUTIONS & TO DUPLICATIONS OR DELETIONS WITHIN OR
CC NEAR DIRECT REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M11077; AAA34285.1; -.
CC DR InterPro; IPR003612; AAI.
CC DR InterPro; IPR001954; Glia_gluutenin.
CC DR PRINTS; PR00208; GLIADGLUTEN.
CC DR SMART; SM00499; AAI; 1.
CC KW Seed storage protein; Repeat; Signal; Multigene family.
CC FT SIGNAL 1 23
CC FT CHAIN 1 304
CC SEQUENCE 304 AA; 34252 MW; 807EBF447A59D6D5 CRC64;
SQ
Query Match 50.7%; Score 38; DB 1; Length 304;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHQR 14
Db 43 SFSQNPVLRKHQR 56

```

```

RESULT      8
ID Y177_TREPA STANDARD: PRT: 436 AA.
AC 063207;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN TP0177.
GN TP0177.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Cholak H., Richardson D., Howell J.K., Chidambaram M., Uitterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RT Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001201; AAC65167.1; -
DR TIGR; TP0177; -.
RW Hypothetical protein: Complete proteome.
SQ SEQUENCE 436 AA; 49532 MW; 50A9DC1D838057AB CRC64;

Query Match 50.7%; Score 38; DB 1; Length 436;
Best Local Similarity 35.7%; Pred. NO. 29;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SFSONPVLKRRHOR 14
   :::::
DB 351 TFSSEGPISRIHKK 364

RESULT      9
ID S0S1_HUMAN STANDARD: PRT: 1333 AA.
AC 007889;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SON OF SEVENTEENLESS PROTEIN HOMOLOG 1 (SOS-1).
GN S0S1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93262494; PubMed=8493579;
RA Chardil P., Camonis J.H., Gale N.W., van Aelst L., Wiegler M.H.,
RA Bar-Sagi D.;
RT "Human SOS1: a guanine nucleotide exchange factor for Ras that binds
RT to GRB2."

```

RL	Science	260:1338-1343(1993).			
RL	[2]				
RP	STRUCTURE BY NMR OF 422-551.				
RX	MEDLINE=96043737; PubMed=9374522;				
RA	Zheng J., Chen R.H., Corblan-garcia S., Cahill S.M., Bar-Sagi D.,				
RA	Cowburn D.;				
RT	"The solution structure of the plectstrin homology domain of human				
RT	SO51. A possible structural role for the sequential association of				
RT	diffuse B cell lymphoma and plectstrin homology domains.";				
RL	J. Biol. Chem. 272:30340-30344(1997).				
RN	[3]				
RP	X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 198-551.				
RX	MEDLINE=99005193; PubMed=9790532;				
RA	Soisson S.M., Milmann A.S., Uy M., Bar-Sagi D., Kuriyan J.;				
RT	"Crystal structure of the Dbl and plectstrin homology domains from				
RT	the human Src of sevenless protein.";				
RL	Cell 95:259-268(1998).				
CC	-1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.				
CC	-1- SUBUNIT: INTERACTS WITH GRB2.				
CC	-1- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY DOMAIN (DH).				
CC	-1- SIMILARITY: CONTAINS 1 PH DOMAIN.				
CC	-1- SIMILARITY: CONTAINS 1 RASGEF DOMAIN.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; L13857; AAA35913.1; .				
DR	PDB; 1AME; 25-FEB-98.				
DR	PDB; 1DBH; 23-DEC-98.				
DR	MIM; 182530; .				
DR	InterPro; IPR002119; Histone_H2A.				
DR	InterPro; IPR000166; Histone_core.				
DR	InterPro; IPR001849; PH.				
DR	InterPro; IPR000651; RasGEFN.				
DR	InterPro; IPR001895; RasGRP_CCC25.				
DR	InterPro; IPR000219; RhogEF.				
DR	pfam; PF00169; PH; 1.				
DR	pfam; PF00617; RasGEFN; 1.				
DR	pfam; PF00618; RasGEFN; 1.				
DR	pfam; PF00621; RhogEF; 1.				
DR	SMART; SM00414; H2A; 1.				
DR	SMART; SM00233; PH; 1.				
DR	SMART; SM00147; RasGEFN; 1.				
DR	SMART; SM00229; RasGEFN; 1.				
DR	SMART; SM00325; RhogEF; 1.				
DR	PROSITE; PS00720; GDS_CDC25; 1.				
DR	PROSITE; PS50003; PH_DOMAIN; 1.				
RW	Guanine-nucleotide releasing factor; 3D-structure.				
FT	DOMAIN 202 443 DH.				
FT	DOMAIN 444 548 PH.				
FT	DOMAIN 777 963 RASGEF.				
FT	DOMAIN 1258 1261 POLY-PRO.				
SO	SEQUENCE 1333 AA; 152463 MW; C6B89CCA11A8DE45 CRC64;				
Query Match		50.7%;	Score 38;	DB 1;	Length 1333;
Best Local Similarity		87.5%;	Pred. No. 96;		
Matches 7; Conservative		0;	Mismatches 1;	Indels 0;	Gaps 0;
OY	5 NPVYAKRH 12				
	11 11111				
DB	944 NPEVYAKRH 951				
RESULT 10					
DATA_WHEAT					
ID DATA_WHEAT	STANDARD.		PRT;	356 AA.	
NC P10385;					

DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLUTENIN, LOW MOLECULAR WEIGHT SUBUNIT PRECURSOR.
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 ON NCBI_TaxID=4565;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_YAMHILL;
 RX MEDLINE=89083577; PubMed=3205747;
 RA Pitts E.G., Rafalski J.A., Hedgcock C.;
 RT "Nucleotide sequence and encoded amino acid sequence of a genomic
 RT gene region for a low molecular weight glutenin.";
 RL Nucleic Acids Res. 16:11376-11376(1988).
 CC -1- FUNCTION: GLUTENIN IS THE HIGH MOLECULAR WEIGHT SEED STORAGE
 CC PROTEIN OF WHEAT ENDOSPERM. THOUGHT TO BE RESPONSIBLE FOR THE
 CC VISCID-ELASTIC PROPERTY OF WHEAT DOUGH.
 CC -1- SUBUNIT: DISULFIDE-BRIDGE LINKED AGGREGATES.
 CC -1- MISCELLANEOUS: GLUTENINS ARE CODED BY SEVERAL GENES ON EACH OF THE
 CC GROUP 1 CHROMOSOMES OF WHEAT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X07747; CA30570.1; -
 DR PIR: S01992; S01992.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001954; G1a-glutenin.
 DR PRINTS: PR00208; GLIADGLUTEN.
 DR SMART: SM00459; AAI; 1.
 DR Seed storage protein; Repeat; Multigene family; Signal.
 KM SIGNAL 1 356
 FT CHAIN ? 356 GLUTENIN, LOW MOLECULAR WEIGHT SUBUNIT.
 SQ SEQUENCE 356 AA; 41020 MW; AA2613FCDD4DC45 CRC64;
 Query Match 49.3%; Score 37; DB 1; Length 356;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 FSONPYLKRHQ 13
 Db 71 FSOQPTISQOQ 82
 RESULT 11
 ID F16B_MOUSE STANDARD; PRT; 94 AA.
 AC P97323;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE FRUCTOSE-1,6-BISPHOSPHATASE ISOZYME 3 (EC 3.1.3.11) (D-FRUCTOSE-1,6-
 DE BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE) (FRAGMENT).
 GN FBP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=97260614; PubMed=9106734;
 RA Cloix J.F., Beaulieu E., Heyor T.K.;
 RT "Various fructose-1,6-bisphosphatase mRNAs in mouse brain, liver,

RT kidney and heart.";
 RL NeuroReport 8:617-622(1997).
 CC -1- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE + H(2)O -
 CC D-FRUCTOSE 6-PHOSPHATE + ORTHOPHOSPHATE.
 CC -1- PATHWAY: NECESSARY FOR, AND SPECIFIC TO, GLUCONEOGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE FBPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y11067; CA71946.1; -
 DR HSRP: P09467; IFTN.
 DR MGD; MGI:1330236; FBP3.
 DR InterPro: IPR000146; In_FB_phptase.
 DR Pfam: PF00316; FBPase; 1.
 DR ProDom: PD001491; In_FB_phptase; 1.
 DR PROSITE: PS00124; FBPASE; PARTIAL.
 KM Hydrolyase; Carbohydrate metabolism; Gluconeogenesis.
 FT NON_TER 1 94
 FT MON_TER 1 94
 SQ SEQUENCE 94 AA; 10459 MW; CB8987281DB2944 CRC64;
 Query Match 48.0%; Score 36; DB 1; Length 94;
 Best Local Similarity 60.0%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 5 NPPVLKRHR 14
 Db 68 NPALLEYHR 77
 RESULT 12
 ID PROC_NEUCR STANDARD; PRT; 311 AA.
 AC Q12641;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
 GN PRO-1.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 ON NCBI_TaxID=5141;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96004701; PubMed=7565596;
 RA Davis C.R., McPeck M.A., McClung C.R.;
 RT "Molecular characterization of the proline-1 (pro-1) locus of
 RT Neurospora crassa, which encodes delta 1-pyrroline-5-carboxylate
 RT reductase.";
 RL Mol. Gen. Genet. 248:341-350(1995).
 CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-
 CC CARBOXYLATE + NAD(P)H.
 CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U30317; AAA83568.1; -

DR InterPro: IPR000304; P5CR.
 DR Pfam: PF01089; P5CR.1.
 DR PROSITE: PS00521; P5CR.1.
 KW Oxidoreductase; Proline biosynthesis; NADP.
 SO SEQUENCE 311 AA; 32151 MW; 9A0D525D845984DF CRC64;

Query Match 48.0%; Score 36; DB 1; Length 311;
 Best Local Similarity 75.0%; Pred. NO. 45;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 NEPVLRH 12
 ||| |||
 Db 26 NEPALPRH 33

RESULT 13

HTPX_AERPE STANDARD; PRT; 366 AA.

AC Q9YD67;

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE PROBABLE PROTEASE HTPX HOMOLOG (EC 3.4.24.-)

GN HTPX OR APE1045.

OS Aeropyrum pernix.

OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;

OX NCBI_TaxID=56636;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K1;

RX MEDLINE=99310339; PubMed=10382966;

RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y.,

Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankei A., Kosugi H.,

Hosoya A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,

Takamaya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,

Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,

Nakamura Y., Nomura N., Sako Y., Kikuchi H.,

"Complete genome sequence of an aerobic hyper-thermophilic

crenarchaeon, Aeropyrum pernix K1."

RT DNA Res. 6:83-101(1999).

RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M48 (ZINC

METALLOPROTEASE).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC EMBL: AP000060; BA80030.1; -

DR MEROPS: M48.004; -

DR InterPro: IPR001915; Peptidase_M48.

DR InterPro: IPR000130; Zn_MTPeptide.

DR Pfam: PF01435; Peptidase_M48.1.

DR PROSITE: PS00142; ZINC PROTEASE; FALSE NEG.

KW Transmembrane; Hydrolase; Metalloprotease; Zinc; Complete proteome.

FT TRANSMEM 4 24 POTENTIAL.

FT TRANSMEM 50 70 POTENTIAL.

FT TRANSMEM 79 99 POTENTIAL.

FT TRANSMEM 193 213 POTENTIAL.

FT TRANSMEM 229 249 POTENTIAL.

FT METAL 181 181 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 182 182 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 185 185 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 366 AA; 39742 MW; 3DF2AF596CC76584 CRC64;

SO SEQUENCE

Query Match 48.0%; Score 36; DB 1; Length 366;

Best Local Similarity 70.0%; Pred. NO. 53;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 FSONPVLRH 11
 ||: ||| ||
 Db 340 FSTHPVSKR 349

RESULT 14

BMP2_MOUSE STANDARD; PRT; 394 AA.

AC P21274;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE BONE MORPHOGENETIC PROTEIN 2 PRECURSOR (BMP-2) (BMP-2A).

GN BMP2 OR BMP-2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94289485; PubMed=8018727;

RA Feng J.O., Harris M.A., Ghosh-Choudhury N., Feng M., Mundy G.R.,

Harris S.E.,

"Structure and sequence of mouse bone morphogenetic protein-2 gene

(BMP-2): comparison of the structures and promoter regions of BMP-2

and BMP-4 genes."

RT Biochim. Biophys. Acta 1218:221-224(1994).

RL [2]

RP SEQUENCE OF 1-351 FROM N.A.

RX MEDLINE=90228966; PubMed=1970330;

RA Dickinson M.E., Kodrin W.S., Sillan C.M., Kingsley D.M., Justice M.J.,

Miller D.A., Ceci J.D., Lock L.F., Lee A., Buchberg A.M.,

Stracusa L.D., Lyons K.M., Derynck R., Hogan B.L.M., Copeland N.G.,

Jenkins N.A.,

"Chromosomal localization of seven members of the murine TGF-beta

superfamily suggests close linkage to several morphogenetic mutant

loci."

RT Genomics 6:505-520(1990).

RL -1- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.

CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.

CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC EMBL: L25602; AAB05665.1; -

DR PIR: A34201; A34201.

DR HSP: P18075; BMP.

DR MGD: MGI:88177; BMP2.

DR InterPro: IPR001839; TGF-beta.

DR InterPro: IPR001111; TGF-beta.

DR Pfam: PF00019; TGF-beta.1.

DR Pfam: PF00688; TGF-beta.1.

DR Prodom: PD000357; TGF-beta.1.

DR SMART: SM00204; TGF-beta.1.

DR PROSITE: PS00250; TGF-BETA.1; 1.

KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 280

FT CHAIN 281 394

FT DISULFD 294 359

FT DISULFD 323 391

FT DISULFD 327 393

FT DISULFD 358 358

FT CARBOHYD 134 134

BONE MORPHOGENETIC PROTEIN 2.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

INTERCHAIN (BY SIMILARITY).

N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match
Best Local Similarity 48.0%; Score 36; DB 1; Length 394;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

FT	CARBOHYD	162	162	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	198	198	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	336	336	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	110	110	T -> S (IN REF. 2).
FT	CONFLICT	113	114	OL -> HE (IN REF. 2).
FT	CONFLICT	271	271	G -> R (IN REF. 2).
SEQ	SEQUENCE	394 AA;	44514 MW;	FD6A0F10587ED54 CRC64;

Query Match
Best Local Similarity 48.0%; Score 36; DB 1; Length 394;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	4 ONPVLKRHR 14
DB	233 ENPGVSKRHVR 243

Query Match
Best Local Similarity 48.0%; Score 36; DB 1; Length 561;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	6 PVLKRHR 14
DB	551 PKTKRHR 559

Search completed: January 17, 2002, 12:37:52
Job time: 112 sec

RESULT 15
YGG4_YEAST STANDARD; PRT; 561 AA.
ID YGG4_YEAST
AC P53166;
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE PUTATIVE ATP-DEPENDENT RNA HELICASE YGL064C.
GN YGL064C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE OF 1-307 FROM N.A.
RC STRAIN-S288C;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
chromosome VII.";
RL Yeast 13:1077-1090(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE=97377993; PubMed=9234674;
RA Feuerhahn M., de Montigny J., Potier S., Souciet J.-L.;
RT "The characterization of two new clusters of duplicated genes
suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae
chromosomes.";
RL Yeast 13:861-869(1997).
CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z72586; CAA96767.1; -;
CC SGD: S0003032; YGL064C.
CC InterPro: IPR001410; DEAD.
CC InterPro: IPR000629; DEAD_ATP_helase.
CC InterPro: IPR001650; Helicase_C.
CC Pfam: PF00270; DEAD_1.
CC Pfam: PF00271; Helicase_C; 1.
CC SMART: SM00487; DEXdc; 1.
CC SMART: SM00490; HELIC; 1.
CC PROSITE: PS00039; DEAD_ATP_HELICASE; FALSE_NEG.
CC Hypothetical protein; RNA-binding; Helicase.
FT NP_BIND 164 171 ATP (POTENTIAL).
FT SITE 287 290 DEAD BOX.
FT SEQUENCE 561 AA; 63057 MW; A8CEB92D461DB6EC CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:36:00 ; Search time 22.94 Seconds
(without alignments)
89.268 Million cell updates/sec

Title: US-09-740-211-15
Perfect score: 1 SFSQNPVYLRKHQR 14
Sequence:

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 segs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	72.0	2343	6	062730
2	54	72.0	2343	6	018806
3	42	56.0	1077	3	005549
4	41	54.7	188	2	09A4F6
5	41	54.7	450	10	09FMA6
6	40	53.3	330	4	075319
7	40	53.3	330	4	09BWE3
8	40	53.3	423	3	059697
9	39	52.0	225	5	09R3V4
10	39	52.0	291	10	09F1W3
11	39	52.0	445	5	09W0F6
12	39	52.0	608	10	09FHY4
13	39	52.0	697	2	005161
14	39	52.0	3085	12	070710
15	39	52.0	3335	2	09PKM7
16	38	50.7	185	5	09GVT0
17	38	50.7	217	11	09D9Z3
18	38	50.7	269	1	074066
19	38	50.7	273	1	074043

20	38	50.7	303	10	P93792	P93792 tritlicum ae
21	38	50.7	332	2	Q9HV64	Q9HV64 pseudomonas
22	38	50.7	348	10	Q42956	Q42956 nicotiana t
23	38	50.7	604	3	013845	013845 schizosacch
24	38	50.7	623	5	018013	018013 caenorhabdi
25	38	50.7	636	2	09F620	09F620 treponema h
26	38	50.7	874	5	Q9N4G4	Q9N4G4 caenorhabdi
27	38	50.7	1144	5	Q9XW61	Q9XW61 caenorhabdi
28	38	50.7	1217	5	P91094	P91094 caenorhabdi
29	38	50.7	1236	4	Q9C012	Q9C012 homo sapien
30	38	50.7	1292	10	Q9L650	Q9L650 oryza sativ
31	38	50.7	1475	5	Q9G6V1	Q9G6V1 caenorhabdi
32	37	49.3	147	12	Q9QP21	Q9QP21 human limun
33	37	49.3	190	3	Q9UW73	Q9UW73 schizosacch
34	37	49.3	258	2	Q9X4J0	Q9X4J0 streptomyce
35	37	49.3	359	10	P93794	P93794 tritlicum ae
36	37	49.3	360	4	Q9P2Y6	Q9P2Y6 homo sapien
37	37	49.3	373	10	Q9Z1I6	Q9Z1I6 tritlicum ae
38	37	49.3	387	10	Q9FE02	Q9FE02 tritlicum tu
39	37	49.3	432	10	Q9FZ18	Q9FZ18 arabidopsis
40	37	49.3	440	10	Q9FEN3	Q9FEN3 arabidopsis
41	37	49.3	442	10	Q9PED7	Q9PED7 oryza sativ
42	37	49.3	481	10	Q9LZR2	Q9LZR2 arabidopsis
43	37	49.3	511	10	Q9LVB1	Q9LVB1 arabidopsis
44	37	49.3	557	2	Q9ADF7	Q9ADF7 streptomyce
45	37	49.3	624	5	Q9VXU2	Q9VXU2 drosophila

ALIGNMENTS

RESULT 1
ID 062730 PRELIMINARY; PRT: 2343 AA.

AC 062730;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FACTOR VIII
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY, SPLEEN;
RA Gordy P.W., Bowen R.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF049489; AAC05384.1; -
DR HSSP; P00451; ICFG.
DR InterPro: IPR001117; Cu-oxidase.
DR InterPro: IPR000421; FA58_C.
DR Pfam; PF00394; Cu-oxidase; 3.
DR Pfam; PF00754; F5_P8_Type-C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADADDD99 CRC64;

Query Match Best Local Similarity 72.0%; Score 54; DB 6; Length 2343;
Matches 10; Conservative 0; Pred. No. 0.7; Mismatches 2; Indels 0; Gaps 0;

QY 3 SONPVLKRRHQR 14
|||||
Db 1648 SONPVLKRRHQR 1659

RESULT 2
ID 018806 PRELIMINARY; PRT: 2343 AA.

AC 018806;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE FACTOR VIII.
 GN F8.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER.
 RA Cameron C., Noley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
 RA Giles A., Lillietrap D.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF016234; AAB87412.1; -.
 DR HSP: P00451; 1CFG.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR000421; FA58_C.
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR SMART: SM00231; FA58C_2.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 DR SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;
 SQ

Query Match 72.0%; Score 54; DB 6; Length 2343;
 Best Local Similarity 83.3%; Pred. No. 0.7;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 SONPVLRKHQR 14
 ||||| |
 Db 1648 SONPVSKHQR 1659

RESULT 3
 ID 005549 PRELIMINARY; PRT; 1077 AA.
 AC 005549;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE SIMILAR TO SEVERAL DNA HELICASES.
 GN YDR291W OR D9819.1.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C (AB972);
 RA Fulton L.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C (AB972);
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Pavello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
 RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
 RA Johnston Y., Langston Y., Latreille P., Le T., Mardis E., Meneses S.,
 RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,
 RA Taich A., Trevasik E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
 RA Wilson R., Waterston R.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C (AB972);
 RA Waterston R.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.
 RC STRAIN=S288C (AB972);
 RA Tia Y., Cherry J.M.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO DEAD/DEAH BOX HELICASE FAMILY.
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
 DR EMBL: U51031; AAB64466.1; -.
 DR SGD: S0002699; YDR291W.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 1077 AA; 123548 MW; 948F024154FBE9A6 CRC64;

Query Match 56.0%; Score 42; DB 3; Length 1077;
 Best Local Similarity 60.0%; Pred. No. 45;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPPVLRKHQR 14
 ||||| |
 Db 500 NPPVLRKHQR 509

RESULT 4
 ID 09A9F6 PRELIMINARY; PRT; 188 AA.
 AC 09A9F6;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN CC0645.
 GN CC0645.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21173698; PubMed=11259647;
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Pirochka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.C., Durkin A.S., Gwin M.L., Hart D.H.,
 RA Kolony J.F., Sait J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uteberck T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AB005740; AAK22630.1; -.
 DR TIGR: CC0645; -.
 DR KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 188 AA; 21013 MW; D41A5973FRC3B31 CRC64;

Query Match 54.7%; Score 41; DB 2; Length 188;
 Best Local Similarity 54.5%; Pred. No. 12;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 SONPVLRKHQR 13
 ::|||::|
 Db 143 AEGPVLRKHQR 153

RESULT 5
 ID 09F9N6 PRELIMINARY; PRT; 450 AA.
 AC 09F9N6;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.
 OS Arabidopsis thaliana (Mouse ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RX MEDLINE=98162728; PubMed=9501997;
 RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
 RT Sequence features of the regions of 1,191,918 bp covered by seventeen
 RT physically assigned P1 clones."
 RL DNA Res. 4:401-414(1997).
 DR EMBL: AB008264; BAB09184.1; -;
 DR InterPro: IPR003480; Transferase.
 DR Pfam: PF02458; Transferase; 1.
 KW Transferase.
 SQ SEQUENCE 450 AA; 50523 MW; 2325CF8739125A2 CRC64;

Query Match 54.7%; Score 41; DB 10; Length 450;
 Best Local Similarity 77.8%; Pred. No. 29;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 ONPVLKRH 12
 Db 201 KNPPYLKRY 209

RESULT 6
 ID 075319 PRELIMINARY; PRT; 330 AA.
 AC 075319;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PIR1 (EC 3.1.3.48).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98352073; PubMed=9685386;
 RA Yuan Y., Li D.M., Sun H.;
 RT "PIRL, a novel phosphatase that exhibits high affinity to RNA.
 RT ribonucleoprotein complexes."
 RL J. Biol. Chem. 273:20347-20353(1998).
 DR EMBL: AF023917; AAC99925.1; -;
 DR InterPro: IPR000340; DS-phosphatase.
 DR InterPro: IPR000387; TYR-phosphatase.
 DR Pfam: PF00782; DSPC; 1.
 DR SMART: SM00012; PTPC_DSPC; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
 KW Hydrolyase.
 SQ SEQUENCE 330 AA; 38939 MW; 0C397FA3043B450A CRC64;

Query Match 53.3%; Score 40; DB 4; Length 330;
 Best Local Similarity 63.6%; Pred. No. 32;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 FSONPVLKRH 12
 Db 267 FSENPHYQRH 277

RESULT 7

09BME3
 ID 09BME3 PRELIMINARY; PRT; 330 AA.
 AC 09BME3;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE DUAL SPECIFICITY PHOSPHATASE 11 (RNA/RNP COMPLEX 1-INTERACTING).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RHADOMYOSARCOMA;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000346; AAH00346.1; -;
 SQ SEQUENCE 330 AA; 38896 MW; D69FB85F6984FBF8 CRC64;

Query Match 53.3%; Score 40; DB 4; Length 330;
 Best Local Similarity 77.8%; Pred. No. 32;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 FSONPVLKRH 12
 Db 267 FSENPHYQRH 277

RESULT 8
 ID 059697 PRELIMINARY; PRT; 423 AA.
 AC 059697;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PROBABLE SERINE/THREONINE-PROTEIN KINASE C32C12.03C (EC 2.7.1.-).
 CN SPC32C12.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wood V., Rajandream M.A., Barrell B.G., Badcock K., Churcher C.M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: STRONG, TO S.POMBE KIN1.
 DR EMBL: AL022172; CA18163.1; -;
 DR InterPro: IPR000719; Euk-kinase.
 DR InterPro: IPR002290; Ser_thr_kin-actsite.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 53 305 PROTEIN KINASE.
 FT NE_BIND 59 67 ATP (BY SIMILARITY).
 FT BINDING 82 82 ATP (BY SIMILARITY).
 FT ACT_SITE 175 175 BY SIMILARITY.
 SQ SEQUENCE 423 AA; 47673 MW; FDDDB5BF32A76E80 CRC64;

Query Match 53.3%; Score 40; DB 3; Length 423;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 PVLKRHR 14
 Db 393 PVPYSRHR 401

RESULT 9
 Q9W0F6 PRELIMINARY; PRT; 225 AA.
 ID Q9W0F6
 AC Q9W0F6
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE TETRASPANIN
 GN TSP42BA OR BCDNA:GH05668 OR CG18817.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Page-McCaw A.W., Tsang G., Rubin G.M.;
 RT "Sequencing Drosophila cDNAs related to tetraspanins."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DE EMBL: AF220041; AF23825.1;
 DR FLYBase; FBgn0029508; TSP42BA.
 DR InterPro; IPR000301; Transmem_4.
 DR Pfam; PF00335; Transmembrane4; 1.
 DR PRINTS; PR00259; TMFOUR.
 SQ SEQUENCE 225 AA; 25897 MW; 806206990642481D CRC64;

Query Match 52.0%; Score 39; DB 5; Length 225;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 6 PPVLRKHOR 14
 DB 157 PPLLRQHQ 165
 RESULT 10
 Q9W0F6 PRELIMINARY; PRT; 291 AA.
 ID Q9W0F6
 AC Q9W0F6
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE EMBL: CAB88415.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA MEDLINE=59156233; Pubmed=10048488;
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:379-391(1998).
 DE EMBL: AB016877; BAB1631.1;
 DR SEQUENCE 291 AA; 32563 MW; 5F9BE6F48C37F5CD CRC64;

Query Match 52.0%; Score 39; DB 10; Length 291;
 Best Local Similarity 66.7%; Pred. No. 43;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 4 QNPVLRKH 12
 DB 75 ENPPVLRKH 83

RESULT 11
 Q9W0F6 PRELIMINARY; PRT; 445 AA.
 ID Q9W0F6
 AC Q9W0F6
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE CG9169 PROTEIN.
 GN CG9169
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE=20196006; Pubmed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Pacled J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spalding A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhou M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003471; AAF47492.1;
 DR FLYBase; FBgn0035217; CG9169.
 SQ SEQUENCE 445 AA; 51459 MW; 3F6CB46D7424712E CRC64;

Query Match 52.0%; Score 39; DB 5; Length 445;
 Best Local Similarity 42.9%; Pred. No. 65;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SFSQNPVLRKHOR 14
 DB 16 SFRNPDLQHQHQ 29
 RESULT 12
 Q9PHY4

```

ID 09PHY4 PRELIMINARY; PRT; 608 AA.
AC 09PHY4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MAP KINASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=99397451; PubMed=10470850;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT Pl and TAC clones."
RL DNA Res. 6:183-195(1999)
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AB017067; BAB0432.1; -
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TYKc; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 608 AA; 69479 MW; C2AB8C51693CFD5 CRC64;

Query Match
Best Local Similarity 52.0%; Score 39; DB 10; Length 608;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 SONPVLKR 11
| | | | | | |
DB 193 SONPVIHR 201

RESULT 13
Q05161 PRELIMINARY; PRT; 697 AA.
AC 005161;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE O-ANTIGEN BIOSYNTHESIS PROTEIN B.
GN RFB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O101 / B41.
RX MEDLINE=93138438; PubMed=8423009;
RA Cheah K.-C., Manning P.A.;
RT "Inactivation of the Escherichia coli B41 (O101:99/P41) rfb gene
RT encoding an 80-kDa polypeptide results in the synthesis of an
RT antigenically altered lipopolysaccharide in E. coli K-12."
RL Gene 123:9-15(1993).
CC -1- FUNCTION: MAY BE INVOLVED IN MODIFYING O-ANTIGEN CHAIN LENGTH
CC AND/OR LINKAGES BETWEEN THE SUBUNITS.
CC -1- PATHWAY: LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
DR EMBL: X59852; CAA42515.1; -
DR InterPro: IPR001296; Glycos_transf_1.

```

```

DR InterPro: IPR001173; Glycos_transf_2.
DR Pfam: PF00534; Glycos_transf_1; 1.
DR Pfam: PF00535; Glycos_transf_2; 1.
KW Lipopolysaccharide biosynthesis; Transmembrane; Inner membrane.
SQ SEQUENCE 697 AA; 78852 MW; A55ED6FC20ED691 CRC64;

Query Match
Best Local Similarity 52.0%; Score 39; DB 2; Length 697;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 FSONPVLKRHOR 14
| | | | | | |
DB 684 FGOVPEMLKKHOR 696

RESULT 14
Q070710 PRELIMINARY; PRT; 3085 AA.
AC 070710;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE POLYPROTEIN.
OS Infectious flacherie virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae.
OX NCBI_TaxID=12742;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98166871; PubMed=9505971;
RA Isawa H., Asano S., Sanara K., Iizuka T., Bando H.;
RT "Analysis of genetic information of an insect picorna-like virus,
RT infectious flacherie virus of silkworm: evidence for evolutionary
RT relationships among insect, mammalian and plant picorna(-like)
RT viruses."
RL Arch. Virol. 143:127-143(1998).
DR EMBL: AB000906; BAA25371.1; -
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polyprotein.
SQ SEQUENCE 3085 AA; 345800 MW; 23FC9600DBF71531 CRC64;

Query Match
Best Local Similarity 52.0%; Score 39; DB 12; Length 3085;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FSONPVLKRH 12
| | | | | | |
DB 2070 FNONSPWIKOH 2080

RESULT 15
Q09PKM7 PRELIMINARY; PRT; 3335 AA.
AC 09PKM7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ADHERENCE FACTOR.
GN TC0438.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MORN / NIGG.
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heideberg J.F.,
RA White O., Hickey E.K., Peterson J., Utechtack T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,

```

RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AK39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AE002311; AAF39292.1; -.
DR TIGR; TC0438; -.
DR InterPro: IPR001917; AminoTransf_2.
DR PROSITE: PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
KM Complete proteome.
SQ SEQUENCE 3335 AA; 376515 MW; F9F7FD7829CB258 CRC64;

Query Match 52.0%; Score 39; DB 2; Length 3335;
Best Local Similarity 87.5%; Pred. No. 4.7e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 SONPVVK 10
|||||:
Db 2174 SONPPLK 2181

Search completed: January 17, 2002, 12:37:36
Job time: 96 sec